Functional analysis of low-moderate risk susceptibility genes emerging from pathway based approaches and Genome Wide Association Studies in epithelial ovarian cancer

by

Maria Notaridou

University College London

PhD Degree

Declaration

I, Maria Notaridou, confirm that the work presented in this thesis is my own. Work that has derived from other sources is indicated and appropriate references and credits have been included.

Abstract

Pathway based and genome wide association studies aim to identify alleles of low/moderate risk that may account for the development of EOC not attributed to high risk susceptibility genes.

The aim of this study was to investigate the functional role of low/moderate susceptibility SNPs and candidate genes that emerge from candidate gene approaches and GWAS using tumour tissues or appropriate models from the proposed cells of origin for EOC, normal Ovarian Surface Epithelium (NOSE), Fallopian Tube Epithelium (FTE) respectively. Part of this study focused in establishing NOSE, FTE and EOC cell lines to study differential expression of candidate genes in post-GWAS functional characterisation studies. Additionally, I have established 3D FTE cultures and propose they more closely resemble the *in vivo* characteristics than 2D cultures.

A candidate gene approach has identified nine candidate genes. I tested 301 invasive EOC tumours and found frequent LOH for tagging SNPs in those genes. LOH was associated with worse survival for *AXIN2, CASP5, RRBBP8* and *AIFM2* but the result for *AXIN2, CASP5* and *AIFM2* were associated with stage. Additionally, one SNP in *STAG3* showed significant preferential loss of the common allele.

Six loci containing susceptibility SNPs were identified in an ovarian cancer GWAS. I found compelling evidence for the somatic role of several genes within these loci in EOC development based on their differential expression between normal (NOSE & FTE) and EOC cell lines. These genes included *PVT1*, *SP2*, *CBX1*, *PNPO*, *HAUS8*, *USE1*, *SKAP1*, *MERIT40* and members of the *HOXB* family of genes with an observed gain of function role and *TIPARP* and *BNC2* with an observed loss of function role in EOC development. Additionally, I found weak associations of susceptibility SNPs' genotypes with *CBX1, SNX11, SP2 and HOXD1 expression* and compelling evidence of genotype specific methylation of *HOXB5* using non tumour samples.

I further knocked down *MERIT40* in EOC cell lines to study the potential role of the gene in EOC development. I found that MERIT40 depletion led to increased accumulation of spontaneous DNA damage and cell cycle arrest characterised by reduction in ploidy of the mucinous EOC cell line EFO27.

This study provides functional evidence that GWAS are powerful tools for identifying novel genes implicated with EOC and that further functional investigation of GWAS identified loci leads to a better understanding of the molecular players involved in EOC initiation, development and survival.

Acknowledgements

I would like to thank my primary supervisor Dr Susan Ramus for her supervision, help and guidance, without which this project would not be feasible. I am also grateful for her help with proof reading and commenting on this thesis. Additionally, I would like to thank my secondary supervisor Prof Simon Gayther for his quidance and the opportunities he gave me to participate in many exciting projects within our group. I also would like to thank all the members of the Ovarian Cancer Association Consortium and members of the UKOPS study.

I would like to thank several members of my group for their contributions towards this study. First of all, our dedicated and wonderful research assistant, Miss Eva Wozniak, for the help in primary tissue collection and for running the Fluidigm chips and for her valuable help in the organisation of the laboratory. I would like to thank Dr Kate Lawrenson for sharing data that were relevant to my research and for the collaboration for publication of the fallopian tube project. I am grateful to Dr Christopher Jones for the aCGH data but mostly for the endless scientific conversations we had and for his patience to introduce me to statistical analysis in R. I would also like to thank Dr Lydia Quaye and Dr Dimitra Dafou for their contributions in our collaboration on the MMCT-18 project. I would also like to thank Prof Martin Widschwendter for his help by approving collection of cisplatin and carboplatin from UCLH and for making the methylation data available that contributed to a large part of analysis performed for this thesis. Finally, I would like to thank all other previous and current members form my group, Sheetal Dyal, Tanya Levi, Barbara Grun, Kate Thornton, Alison Jones, Mark Cox, and our IT support Vijay Devineni.

I would also like to give special thanks to Dr Liz Benjamin for her valuable input in histopathological examination of several of my specimens. Additionally, I thank all the members of the Gynaecological oncology surgical team for allowing me in surgery endless times until I could collect all the specimens I needed and Miss Nyalla Balogun for consenting the patients.

Finally, I would like to thank my partner and friends for all their patience, support and encouragement along the way.

Dedication

This thesis is dedicated to my wonderful parents, Sofia Ifanti and Kostas Notaridis, who have always supported and believed in me.

Η παρούσα διατριβή είναι αφιερωμένη στους υπέροχους γονείς μου, Σοφία Υφαντή και Κώστα Νοταρίδη, που ποτέ δεν έπαψαν να με στηρίζουν και να πιστεύουν σε μένα. Δεν θα μπορούσα να τα έχω καταφέρει χωρίς εσάς να αγωνίζεστε και να ονειρεύεστε για μενα από τα πρώτα βήματα μου. Σας ευχαριστώ για όλα και σας αγαπώ πολύ.

Table of contents

List of Figures

List of Tables

List of abbreviations

aCGH	:	**a**rray **C**omparative **G**enomic **H**ybridisation
AML	:	**A**cute **M**yeloid **L**eukemia
APC	:	**A**denomatous **P**olyposis **C**oli
AS LOH	:	**A**llele **S**pecific **L**oss **O**f **H**eterozygosity
ASE	:	**A**llele **S**pecific **E**xtension
AS-PCR	:	**A**llele **S**pecific **P**olymerase **C**hain **R**eaction
AS-SBE	:	**A**llele **S**pecific **S**ingle **B**ase **E**xtension
BER	:	**B**ase **E**xcision **R**epair
BPE	:	**B**ovine **P**ituitary **E**xtract
BrDU	:	**Br**omo**D**eoxy**U**ridine
CA125	:	**C**ancer **A**ntigen 125
CCC	:	**C**lear **C**ell **C**arcinoma
cDNA	:	**C**omplementary **DNA**
CFE	:	**C**olony **F**orming **E**fficiency
CGH	:	**C**omparative **G**enomic **H**ybridisation
CI	:	**C**onfidence **I**nterval
CK	:	**C**yto**k**eratin
CN	:	**C**opy **N**umber
CSC	:	**C**ancer **S**tem **C**ells
DDR	:	**D**NA **D**amage **R**esponse
DMSO	:	**D**i**m**ethyl **S**ulf**o**xide
DNA	:	**D**eoxyribo**n**ucleic **A**cid
dNTP	:	**d**eoxyribo**n**ucleotide **T**ri**p**hosphate
DSB	:	**D**ouble **S**trand **B**reaks
DT	:	**D**izygotic **T**wins
EC	:	**E**ndometrioid **C**arcinoma
ECM	:	**E**xtra**c**ellular **M**atrix
EGF	:	**E**pidermal **G**rowth **F**actor
EGRF	:	**E**pidermal **G**rowth **F**actor **R**eceptor
EMT	:	**E**pithelial to **M**esenchymal **T**ransition
EOC	:	**E**pithelial **O**varian **C**ancer
eQTL	:	**e**xpression **Q**uantitative **T**rait **L**oci
ER	:	o**e**strogen **r**eceptor
FACS	:	**F**luorescence-**A**ctivated **C**ell **S**orting
FBS	:	**F**oetal **B**ovine **S**erum
FIGO	:	**I**nternational **F**ederation of **G**ynaecology and **O**bstetrics
FISH	:	**F**luorescence **I**n **S**itu **H**ybridization
FITC	:	**F**luorescein **I**so**t**hio**c**yanate
FSP	:	**F**ibroblast **S**pecific **P**rotein
FT	:	**F**allopian **T**ube

FT	:	**F**allopian **T**ube
FTE	:	**F**allopian **T**ube **E**pithelial
FTSEC	:	**F**allopian **T**ube **Sec**retory
FVIII	:	**F**actor **VIII**
GFP	:	**G**reen **F**luorescent **P**rotein
GGR	:	**G**lobal **G**enomic **R**epair
GST	:	**G**lutathione **S**-**t**ransferases
GTPases	:	**G**uanosine **t**ri-**p**hosphatases
GWAS	:	**G**enome **W**ide **A**ssociation **S**tudies
H&E	:	**H**eamatoxylin and **E**osin
HAPMAP-CEU	:	**Hap**lotype **Map**-**C**aucasian residents of **Eu**ropean ancestry
HBOC	:	**H**ereditary **B**reast/**O**varian **C**ancer **S**yndrome
HC	:	**H**ydro**c**ortisone
HGSC	:	**H**igh **G**rade Serous **C**arcinoma
HIV	:	**H**uman **I**mmunodeficiency **V**irus
HNPCC	:	**H**ereditary **N**on**p**olyposis **C**olorectal **C**ancer
HR	:	**H**omologous **R**ecombination
HRT	:	**H**ormone **R**eplacement **T**herapy
hTERT	:	**h**uman **T**elomerase **R**everse**T**ranscriptase
HUVEC	:	**H**uman **U**mbilical **V**ein **E**ndothelial **C**ells
HWE	:	**H**ardy **W**einberg **E**quilibrium
IC	:	**I**nclusion **C**ysts
IC50	:	**I**nhibitory **C**oncentration half maximal
IOSE	:	**I**mmortilised normal **O**varian **S**urface **E**pithelial
IR	:	**I**onising **I**rradiation
LD	:	**L**inkage **D**isequilibrium
LGSC	:	**L**ow **G**rade **S**erous **C**arcinoma
LOH	:	**L**oss **O**f **H**eterozygosity
MAF	:	**M**inor **A**llele **F**requency
MALOVA	:	**Mal**ignant **Ova**rian cancer prediction
MC	:	**M**ucinous **C**arcinoma
MDA	:	Multiple **D**isplacement amplification
MDR1	:	**M**ulti**d**rug **R**esistance **1**
MET	:	**M**esenchymal to **E**pithelial **T**ransition
miRNA	:	**mi**cro**RNA**
MMCT	:	**M**icrocell-**M**ediated **C**hromosome **T**ransfer
MMR	:	**M**ismatch **R**epair **P**athway
mRNA	:	**m**essenger **RNA**
MS	:	**M**ass **S**pectrometry
MSI	:	**M**icro**s**atellite **I**nstability
MT	:	**M**onozygotic **T**wins
MTT	:	3-(4,5-Dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide
NCBI	:	**N**ational **C**enter for **Bi**otechnology **I**nformation
NER	:	**N**ucleotide **E**xcision **R**epair

NHEJ	:	**N**on-**H**omologous **E**nd **J**oining
NOSE CM	:	**NOSE C**ulturing **M**edia
NOSE	:	**N**ormal Ovarian Surface Epithelial
OC	:	**O**ral **C**ontraceptive
OCAC	:	**O**varian **C**ancer **A**ssociation **C**onsortium
OSE	:	**O**varian **S**urface **E**pithelium or **E**pithelial
OthC	:	**Oth**er **C**arcinoma
PCR	:	**R**olymerase **C**hain **R**eaction
PI	:	**P**ropidium **I**odide
PI3K	:	**P**hospho**i**nisitide **3-k**inase
PR	:	**P**rogesterone **R**eceptor
QC	:	**Q**uality **C**ontrol
RB1	:	**R**etino**b**lastoma 1
RCCS	:	**R**otary **C**ell **C**ulture **S**ystem
RFLP	:	**R**estriction **F**ragment **L**ength **P**olymorphisms
RNA	:	**R**ibo**n**ucleic **A**cid
ROS	:	**R**eactive **O**xygen **S**pecies
RPPA	:	**R**everse **P**hase **P**rotein **A**rray
SC	:	**S**erous **C**arcinoma
shRNA	:	**s**mall **h**airpin **RNA**
SNP	:	**S**ingle **N**ucleotide **P**olymorphisms
SOC	:	**S**erous **O**varian **C**arcinomas
SSB	:	**S**ingle **S**trand **B**reaks
ssDNA	:	**s**ingle **s**trand **DNA**
TAH	:	**T**otal **A**bdominal **H**ysterectomies
TCGA	:	**T**he **C**ancer **G**enome **A**tlas **R**ese**a**rch Network’
TCR	:	**T**ranscription **C**oupled **R**epair
TFU	:	**T**rans**f**orming **U**nits
TGF-β	:	**T**ransforming **G**rowth **F**actor **β**
TIC	:	**T**ubal **I**ntraepithelial **C**arcinoma
TSG	:	**T**umour **S**uppressor **G**ene
TSGA	:	**T**he **C**ancer **G**enome **A**tlas **R**esearch **N**etwork
tSNP	:	**t**agging **SNP**
UCLH	:	**U**niversity **C**ollege **L**ondon **H**ospital
UKOPS	:	**U**nited **K**ingdom **O**varian **P**opulation **S**tudy
UTR	:	**Un**t**r**anslated **R**egion
UV	:	**U**ltra**v**iolet

1 Literature review

1.1 Epithelial Ovarian Cancer: An overview

Epithelial ovarian cancer (EOC) constitutes up to 90% of primary ovarian tumours. It is the most common cause of gynaecological cancer deaths and the fifth most common cancer responsible for female fatalities due to cancer (Shan et al, 2009) with survival rates not improving over the last twenty years (Figure 1.1). In 2002 EOC accounted for 125,000 deaths per year out of 200,000 incidents occurring worldwide (Parkin et al, 2005), increasing to 225,000 incidents and 140,200 deaths in 2008 (Jemal et al, 2011). The risk of incidence for EOC is as high as 1% for the total female population with a 60% risk of mortality for the inflicted population (Jemal et al, 2011).

One of the reasons that the mortality rates for ovarian cancer are so high is that about 70% of the women are diagnosed at advanced stages of the disease (Goff et al, 2007). Early diagnosis can be difficult since the symptoms, in most cases, are not evident in early disease. Several studies have tried to identify a panel of symptoms in early disease that would help early referral of women to a gynaecologist. Most of the reported symptoms are gastrointestinal and urinary symptoms such as persistent abdominal distension, some loss of appetite, pain in the abdomen, fatigue, increased urinary frequency and constipation (Goff et al, 2007, Bankhead et al, 2008). Some studies report that gynaecological symptoms are not evident (Goff et al, 2007), which is contradicted by other studies that have proposed vaginal bleeding independent of menstrual cycle as a symptom (Lurie et al, 2009).

When ovarian cancer is suspected, screening is performed by transvaginal ultrasonography and the measurement of tumour marker cancer antigen CA125 in the blood of the patients is taken (Cannistra et al, 2004). Epithelial ovarian cancer diagnosis is staged during laparotomic procedures and confirmed by biopsy and histopathological examination of the tumour by a pathologist where the grade of the disease is assigned to the tumour according to the International Federation of Gynaecology and Obstetrics (FIGO) (Nam et

al, 2010). Recent studies have shown that a more reliable method for diagnosing malignant EOC and revealing metastatic disease is positron emission tomography/computed tomography (Nam et al, 2009).

According to FIGO there are 4 stages of EOC, Stage I is when the tumour is limited to the ovaries, stage II when there is pelvic extension of the ovarian tumour to structures like the uterus, fallopian tubes or pelvic peritoneum, stage III when the tumour involves the upper abdomen and lymph nodes are involved and Stage IV when there is apparent metastasis in distal organs (Cannistra, 2004). Unfortunately, early stages of EOC rarely exhibit symptoms that will alert the patient to proceed with screening and therefore diagnosis is made mostly at later stages of the disease when the survival rates are significantly reduced (Brun et al, 2000) as documented by Cancer Research UK (Figure 1.1).

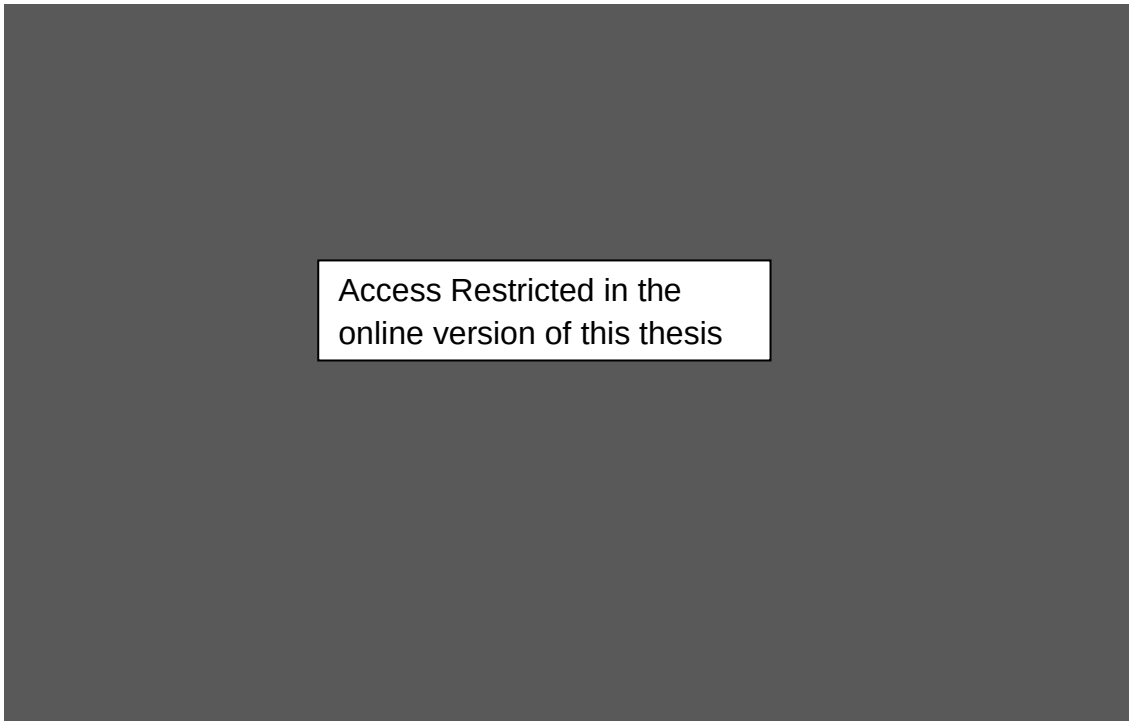


Figure 1.1: Stage of epithelial ovarian cancer associated with survival from disease. Five year stage specific relative survival rates in the UK over the last 20 years as published on March 2011 by Cancer research UK, [Adults (Ages 15-99),Anglia Cancer Network, 1987-2008)]. (http://www.cancerresearchuk.org/cancer-info/cancerstats/types/ovary/survival/ovarian-cancer-survival-statistics).

When EOC is confirmed and depending on the stage at diagnosis the patient would be subjected to total abdominal hysterectomy and/or combination chemotherapy. In early stage disease adjuvant platinum chemotherapy

including cisplatin and carboplatin is used post surgery whereas in advanced stage disease combination of platinum and taxane (the derivative paclitaxel) based chemotherapy is used pre and post surgery (Cannistra, 2004) and possibly radiotherapy as well (http://cancerhelp.cancerresearchuk.org/type/ovarian-cancer/treatment/).

EOC is a complex cancer regarding its initiation and also its metastasis due to the extreme molecular heterogeneity of the disease. Numerous studies have demonstrated several alterations in EOC that involve cell cycle regulation and DNA repair pathways as well as oncogenes and tumour suppressor genes (Darcy et al, 2010). Historically, the etiology for EOC has been studied through clinical observations but due to the lack of early stage specimens our understanding of the initiation and progression of the disease is still poor. The development of appropriate *in vitro* models and animal models is essential to study the underlying causes for EOC development.

1.1.1 Subtypes of epithelial ovarian cancer

EOC should not be treated as a single disease as it is stratified into various histological subtypes. All of the histological subtypes are distinct in their clinical behaviour as well genetic risk factors and molecular pathways implicated in their oncogenesis. Subtypes also have different responses to chemotherapy causing the 5-year survival time to be different for distinct subtypes (Kurian et al, 2005). Reproducible sub-classification of EOC subtypes is very important and although classification of EOCs started as an empirical observation of traditional histomorphologic features, the classification has become more logical by linking these features with expression of particular biomarkers and mRNA profiling and clinical behaviour of the tumours.

The most common histological subtypes of epithelial cancer are serous adenocarcinoma which is subdivided further into high grade serous and low-grade serous, intestinal mucinous, endometrioid and clear-cell. All subtypes are characterised by mutations in different genes and expression of distinct biomarkers (Figure 1.2). Some subtypes are classed as undifferentiated or mixed which are most likely serous subtypes.

Serous carcinomas

These are the most common amongst EOCs accounting for about 80% of all ovarian carcinomas and are evident in about 95% of patients diagnosed with FIGO stages III-IV (Soslow, 2008). The morphology of serous tumours is very heterogeneous which could possibly suggest that they arise from transformation or progression from other types of tumours. Their architecture is predominately papillary containing psammona bodies (Cannistra et al, 2005) mimicking yolk sack tumours but can also be glandular mimicking embryonal carcinomas or solid mimicking dysgerminomas (reviewed in Soslow, 2008).

High grade serous carcinomas (HGSCs) and low grade serous carcinomas (LGSCs) exhibit very similar histologic appearance but as a first step for differential diagnosis, searching for a precursor lesion is indicative, as high grade have been associated with tubal intraepithelial tumours and low grade associated with borderline grade tumours (Soslow, 2008). The discrimination between them is mainly based on some differences on their immunophenotype, genetic profile and their clinical behaviour.

Whereas Wilms tumour 1 (WT1), mesothelin, oestrogen receptor (ER) and cancer antigen 125 (CA125) are commonly over-expressed in all serous carcinomas (SCs), p53 (tumour protein 53) and progesterone receptor (PR) over-expression are uniquely found in the high grade and low grade SCs respectively. Another major difference between the two grades is the significantly higher expression of the E3-ubiquitin-protein ligase Ki-67 (encoded by *Mib-1*) in high grade SCs (Köbel et al, 2008). High grade SCs frequently have *BRCA1*, *BRCA2, TP53 and KRAS* mutations whereas low grade SCs have mutations on *KRAS* and *BRAF* genes (Sieben et al, 2004, Risch et al, 2006, Ahhmed et al, 2010).

Some studies show that SCs have the highest mortality rates between all the EOC subtypes (Brun et al, 2000) with high grade SCs having the poorest prognosis but there is a lot of controversy on that subject as other studies reported SCs to be more responsive to cisplatin treatment together with endometrioid carcinomas (Shimizu et al, 1997). The improved response to cisplatin based chemotherapy is seen especially in patients bearing *BRCA1* and *BRCA2* germline mutations (found in high grade SCs) and thus having better 3-

year survival rates of 60% in mutations carriers compared to 40% in non carriers (Ben et al, 2002). Recent evidence have shown that *BRCA1* and *BRCA2* mutation carriers have improved 5-year survival, 44% and 52% respectively compared to 36% of non-carriers and *BRCA2* mutation carriers have the best prognosis (Bolton et al, 2012).

Intestinal mucinous carcinomas

Primary ovarian mucinous carcinomas (MCs) account for less than 3% of all ovarian carcinomas. About 65% of MCs are FIGO stage I and about 30% of all FIGO stage I patients have MCs. Their morphology is mainly glandular, and presence of intracytoplasmic mucin and goblet cells resembling the lining of the intestine (hence the description intestinal) are essential for a pathologist to classify a tumour as being mucinous. Mucinous borderline and benign tumours also exist and can be distinguished by the absence of goblet cells. Most primary mucinous carcinomas show transitions from cystadenomas to intestinal mucinous borderline to carcinoma (Soslow, 2008).

MCs lack expression of ER, WT1 and CA125 and have high expression of CK7. They often have *KRAS* mutations but they do not possess *BRCA1*, *BRCA2* and *TP53* mutations often seen in SCs (Cuatrecasas et al, 1997).
Mucinous tumours have been reported to belong to a cisplatin resistant group (together with clear cell carcinomas) by demonstrating only a 5% response to cisplatin based chemotherapy (Shimizu et al, 1997). However, Harrison et al have reviewed retrospective studies and show early MCs have a 90% 5-year survival rate whereas only advanced stage MCs present relative resistance to platinum based therapy, with overall response rates of 26-42 % to cisplatin based first line chemotherapy (Harrison et al, 2008).

Endometrioid Carcinomas

Endometrioid carcinomas (ECs) comprise the second most prevalent EOC subtype accounting for about 10% of all EOCs. Most ECs are classified as FIGO stage I and II and they are the most frequent carcinomas observed in FIGO stage I patients constituting about 50% of all stage I cases. Their morphology can be papillary or glandular mimicking Sertoli-Leydig cell tumours,

or solid mimicking granulosa cell tumours. Their architecture resembles the endometrium epithelium and ECs have been associated with endometriosis, endometrioid borderline tumours, or an endometrioid type of synchronous endometrial neoplasms.

Genetic alterations contributing to the development of ECs are *PTEN* mutations (Obata et al, 1998) and mutations in the *CTNNB1* gene encoding for b-catenin (Sagae et al, 1999). The biggest challenge for the pathologists has been to discriminate between SCs and ECs given their molecular and morphological similarity. Studies have shown that high grade ECs lack expression of WT1 and p53 but have high expression of b-catenin whereas about 30% of low grade ECs exhibit p53 over-expression (Madore et al, 2009). Köbel and colleagues have also reported that ECs present high expression of p21, PR and ER.

Endometrioid histology is associated with a better survival compared with SCs, regardless of response to platinum based chemotherapy (Storey et al, 2008) and regardless of the fact ECs have been classified as cisplatin resistant ovarian carcinomas (Shimizu et al, 1997).

Clear cell carcinomas

Clear cell carcinomas (CCCs) account for about 5% of all ovarian tumours. Most CCCs are classified as FIGO stage I and II and they are the second most frequent carcinomas observed in FIGO stage I patients constituting about 35% of all stage I cases. Their architecture is mainly tubulocystic, and solid. Similar to ECs they are associated with an origin of epithelial endometriosis (Soslow, 2008). The molecular events characterising the progression of CCCs are not as well described as for other subtypes. Some studies have reported mutations in the *PTEN* gene (Sato et al, 2000) and *ARID1A*, a gene involved in chromatin remodelling (Wiegand et al, 2010). The latter report showed presence of *ARID1A* mutations in ECs as well. These reports suggest that clear cell tumours could arise from ovarian and pelvic endometriosis.

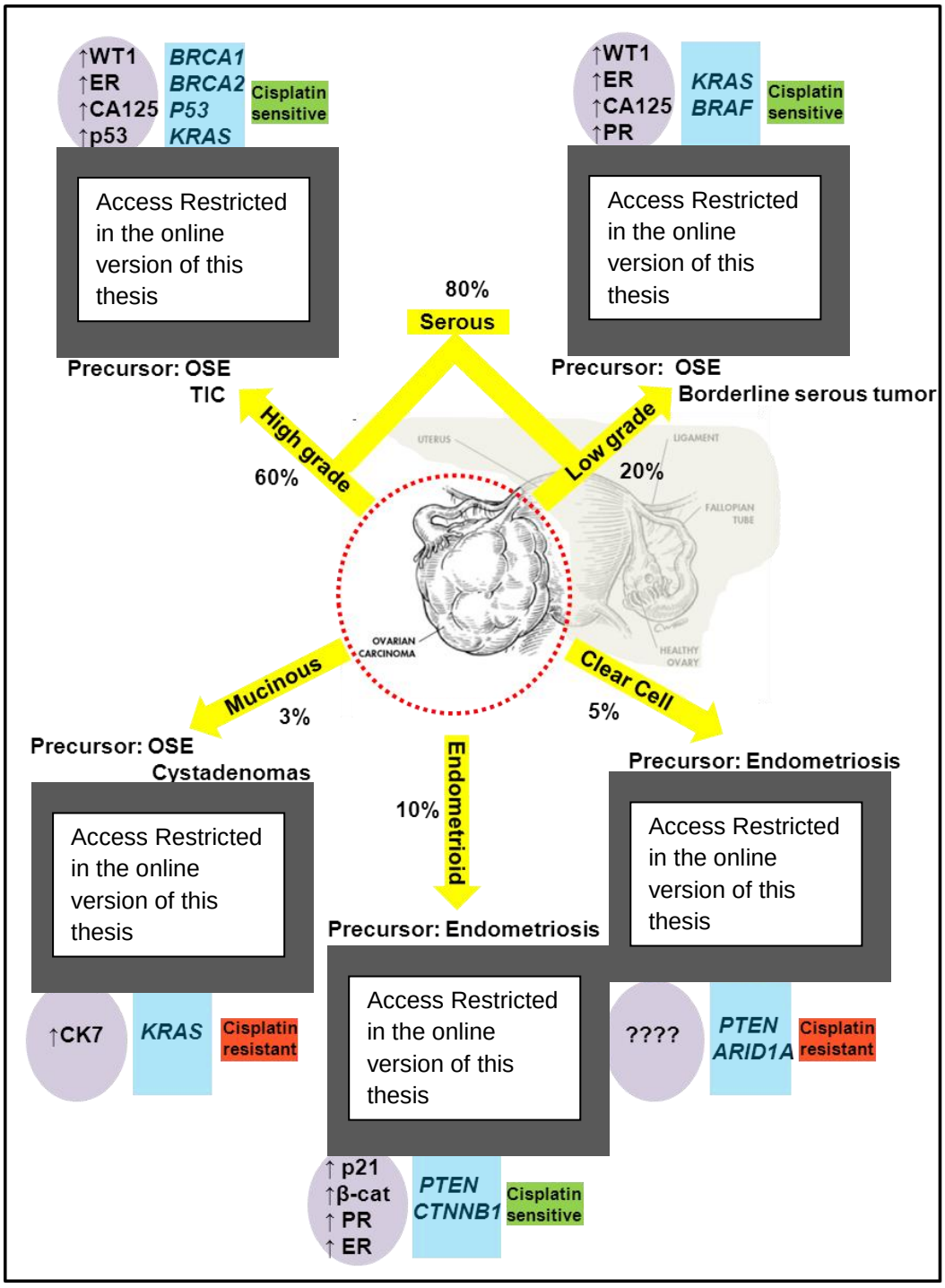


Figure 1.2: Representation of the morphology, proposed precursor lesions, immunophenotype, genetic alterations and clinical behaviour of the five main subtypes of epithelial ovarian cancer. The high grade serous picture shows a papillary pattern and high nuclear grade. The low grade serous picture shows stromal invasion of low grade carcinoma in a borderline serous tumour. The mucinous picture shows mucinous borderline tumour (right part of the picture) continuing to mucinous carcinoma (left part of the picture). The endometrioid picture shows endometrioid adenocarcinoma arising from endometriosis. The clear cell picture shows clear cell carcinoma with microcystic pattern. The pictures of the H&E stained tumours were taken from Soslow, 2008 (high grade serous and clear cell) and Bell, 2005 (low grade, mucinous, endometrioid). TIC: tubal intraepithelial carcinoma. OSE: Ovarian surface epithelium.

1.1.2 Recent models for classification of epithelial ovarian tumours

The histological subtypes of EOC have been more recently categorised into 2 different types, Type I and Type II tumours. This model does not replace the traditional histopathological classification for ovarian carcinoma but provides a framework for the study of ovarian carcinogenesis that is believed to have an important clinical implication as it proposes that the two genetically distinct types of EOC should be differentially treated.

Type I epithelial ovarian tumours comprise of the less frequent low grade and borderline serous tumours as well as endometrioid, mucinous and clear-cell cancers. These histological subtypes are characterised by a distinct mutation spectrum that includes *PTEN, PIK3CA, KRAS, BRAF and CTNNB1* mutations (Kurman et al, 2008, Steffensen et al, 2011). Type II epithelial ovarian tumours comprise of the high-grade serous carcinomas, mixed malignant mesodermal tumours, carcinosarcomas and undifferentiated tumours. The majority of these tumours exhibit high prevalence of *TP53* mutations and loss of *BRCA1* (Kurman et al, 2008, Steffensen et al, 2011). The best serological biomarkers to date for the distinction of type I for type II EOC patients are p53 autoantibodies in combination with CA125 levels, which are absent in the blood of type I patients (Lu et al, 2011).

This classification theory is supported by the dualistic model of serous carcinogenesis according to which the precursors of low-grade serous carcinomas are well characterised precursor lesions called borderline tumours, whereas high-grade-serous carcinomas are genetically distinct and to not represent a transition from low-grade to a high-grade phenotype but arise *de novo* (Singer et al, 2005).

1.2 Risk and protection factors for ovarian cancer

A great proportion of ovarian cancer cases is associated with several risk factors including the incessant ovulation and endometriosis. Risks associated with genetic factors and hereditability (family history) will be also discussed later. Non genetic risk factors include ageing, the use of fertility drugs and hormone replacement therapy (HRT), using talc powder, being overweight, dietary habits and smoking. Protective factors include pregnancy/childbearing, breastfeeding, tubal ligation, use of the oral contraceptive pill and consumption of some painkillers.

1.2.1 Risk Factors for ovarian cancer

Ageing and ovarian cancer: More than 45% of women diagnosed with ovarian cancer are above 65 years old and the older the woman is the highest the risk of being diagnosed with higher stage of the disease (Yancik, 1993). According to the American National Cancer Institute more than 70% US incidence of ovarian cancer over the last 25 years affected women aged 65 or more years old (Figure 1.3).

Use of fertility drugs and ovarian cancer: Infertility has been long recognised as a risk factor for ovarian cancer. Epidemiological studies have reported an increased risk to ovarian cancer amongst infertile women (Risch et al, 1994). Fertility drugs such as human menopausal gonadotropin have been associated with an elevated risk especially in borderline ovarian tumour cases (Harris et al, 1992, Goldberg et al, 1992).

Use of hormone replacement therapy (HRT) and ovarian cancer: There are several types of HRT and include oestrogen and/or progesterone. Several studies, have reported that for women taking HRT more than 5 years the risk of ovarian cancer increases and the longer the consumption is continued the higher the risk (Folsom et al,. 2004, Beral et al, 2007). The most comprehensive study to date, known as the Million Women study showed that since 1991, the use of HRT has contributed to the development of 1300 additional ovarian cancers and 1000 additional deaths in the UK alone (Beral et al, 2007).

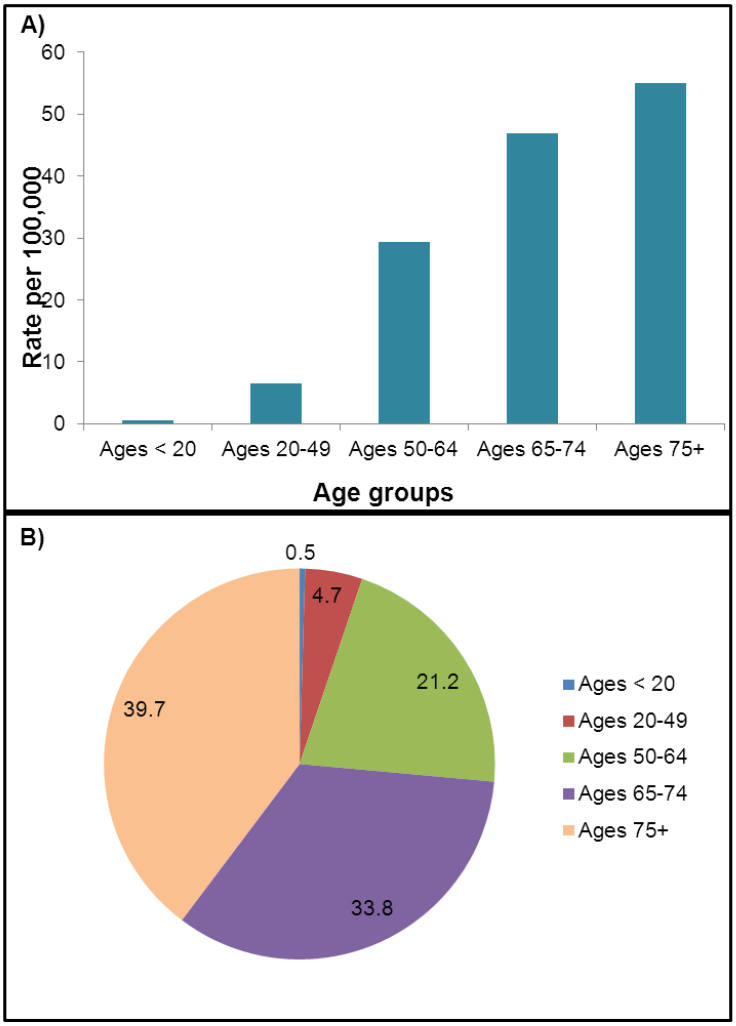


Figure 1.3**:** Epithelial ovarian cancer incidence associated with age. A) Age adjusted rates per 100,000 of population for ovarian cancer in 6 different age groups. The incidence rate is significantly higher in the age groups above 65 years old. B) Percentage of ovarian cancer incidence as distributed between the different age groups. These graphs were designed from data downloaded from the National Cancer Institute and are presenting US incidence from 1992 to 2008 in all ethnic groups as reported by April 2011.

Use of talc powder and ovarian cancer: The association between use of talcum powder genitally or perineally and risk for ovarian cancer has not been consistent amongst studies. Some studies have reported 30% increased risk associated with talcum use (Cramer et al, 1999 May, Cramer et al, 1999 July) or more modest associations (Rosenblatt et al, 2011), but other epidemiological studies did not support a causal association (Huncharek et al, 2011)

Excess weight, diet, smoking and other risk factors and ovarian cancer: There are controversial data showing an association between the use of intra-uterine device and increased risk (Zapata et al, 2010). Conflicting data have also been described for lactose, saturated fat, and also excess weight as ovarian cancer risk factors (Cancer Research UK). Smoking has been associated with increased risk of mucinous cancer in a systematic review of 30 case studies that investigated subtype specific ovarian cancer risk linked with smoking (Jordan et al, 2006) but the data are controversial.

1.2.2 Protective factors for ovarian cancer

Pregnancy: Pregnancy confers protection against ovarian cancer in line with the incessant ovulation hypothesis which will be later discussed. The ovarian surface repair does not take place during pregnancy as there are no ovulatory cycles.

Tubal ligation: It has been shown that women that undergo tubal ligation as a method of sterilisation have a reduced risk in developing ovarian cancer (Tworoger et al, 2007, Cibula et al, 2011). The protective mode of tubal ligation could be attributed either by not allowing ovulation and in turn post-ovulatory repair and also stopping the flow of hormones that make the ovarian microenvironment more mitogenic or by not allowing the cells from endometrium or endosalpinx to transport in the ovary (Dubeau, 2008).

Use of the oral contraceptive (OC) pill: It is well established that use of the OC pill reduces the risk for ovarian cancer. The more prolonged the intake of OC the greater the reduction of the risk for ovarian cancer (Ness et al, 2001, La Vecchia, 2006, Tworoger et al, 2007).

Breastfeeding: A large cohort study has reported a weak association of breastfeeding with reduction in the risk for ovarian cancer compared to women that never breastfed (Danforth et al, 2004) and other studies further supported this evidence.

Consumption of painkillers: There is conflicting data on whether the use of anti-inflammatory drugs such as aspirin or paracetamol confers a protective role against ovarian cancer (Cancer research UK).

1.3 Origins of Epithelial Ovarian Cancer

The genetic, clinical and histological heterogeneity observed in EOC subtypes is likely to be in part due to different cellular origins from which each histotype is derived. The site of origin and the initial events for tumorigenesis in EOC is a question that must be studied to improve earlier detection and better targeted treatments for distinct subtypes.

Ovarian and generally pelvic malignancies are proposed to have three distinct sites of origin:

1. The ovarian surface epithelium (OSE)
2. The Fallopian tube epithelium (FTE)
3. The peritoneal surface

In an attempt to better classify the origins for each EOC subtype an overview through developmental and reproductive biology is essential.

1.3.1 Development of the female reproductive tract and ovulation in adults

Soon after fertilisation of the oocyte with a sperm, a single diploid cell is formed called a zygote. The zygote then undergoes cleavage (cell division) to form a single layer sphere of cells called blastula. Gastrulation is that point of embryonic development where a three layered structure, called gastrula, is formed from the blastula. These three layers are the ectoderm (outer layer), the mesoderm (intermediate layer) and the endoderm (inner layer). The development of the urogenital system, including the reproductive tract, starts soon after the gastrulation event when the mesoderm differentiates. The mesoderm segregates into the pluripotent mesenchyme, which will later give rise to the connective tissues, and the coelomic (peritoneal) mesothelium that gives rise to the müllerian duct. The müllerian mesothelium (müllerian duct), also known as the paramesonephric duct, in turn differentiates into oviductal, endometrial and cervical epithelia (Acien et al, 1992, Murdoch and McDonnel, 2002, Kobayashi et al, 2003, Sajjad, 2010).

The involvement of homeobox genes across the Mullerian duct in the differentiation of the distinct organs has been well established with HOXA9 highly expressed where the fallopian tube will develop, HOXA10 is expressed as the endometrium develops and HOXA11 is expressed in the lower cervix. HOXA genes are also activated in the relevant tissue of the adult female reproductive system (Taylor et al, 1997, Du and Taylor, 2004).

Distinct from this müllerian developmental pathway, the ovarian cortex originates from invagination of the coelomic mesothelium over the gonadal ridge which is the precursor to the gonads (Byskov, 1986). The outer layer of the ovarian cortex consists of a single layered germinal epithelium known as the ovarian surface epithelium (OSE). OSE cells are supported by a basal layer of connective tissue called the tunica albuginea which is enriched with fibroblasts, endothelial, smooth muscle and interstitial cells.

1.3.2 Ovarian Surface Epithelium origin of EOC

A single cell of origin for EOC being the OSE has been widely accepted for several decades but the resemblance of epithelial ovarian tumours to organs and epithelium that derive via Mullerian development raised the question on whether there are additional sites of EOC initiation.

OSE is a single layer of epithelial cells with cuboidal or squamous morphology. They are often described as uncommitted mesothelial cells expressing mesenchymal and epithelial markers due to their shared characteristics with the mesothelial lining of the peritoneum and normal epithelium. OSE cells have been isolated (Auersperg et al 1984) and their culturing conditions have been optimised using a 105:199 media supplemented with foetal bovine serum (FBS), insulin, bovine pituitary extract (BPE), hydrocortisone (HC, Li et al, 2004). OSE immortalisation (Maeda et al 2005) has also been established and optimised over the past decades. The fact that OSE cells have a great phenotypic plasticity has been considered an argument for their potential to differentiate into cancer cells (Connolly et al, 2003).

The OSE has a well-established role in transporting nutrients and mediating the repair of the ovarian epithelium after ovulation. During ovulation the cell-cell adhesions between the OSE cells are broken down by excreted

lysozymes for the mature oocyte to be released. Following ovulation the OSE drastically proliferate at the site of the trauma and repair the surface. This vigorous healing process causes the formation of clefts in the cortical stroma. These clefts have the tendency to close up and form invaginations at the cortical ovarian stroma, and are then called cortical inclusion cysts (ICs). Within the ovarian stroma, the OSE-lined inclusion cysts are exposed to a more mitogenic environment that might be promoting tumour development (Burdette et al, 2007, Risch, 2007).

There are three main hypotheses that support the OSE origin of ovarian carcinogenesis and there is a large amount of molecular evidence that support these theories:

A) The incessant ovulation hypothesis

The first one to suggest this hypothesis was Fathalla et al in 1971. She performed an experiment where hens were subjected in hyperovulation by exposing them to prolonged photostimulation and reported the development of ovarian adenocarcinomas in 17 out of 19 experimental subjects with the control hens remaining healthy (Fathalla, 1971). This hypothesis proposes that the higher the number of ovulation cycles in a lifetime of a female the higher the chances are for the repair mechanisms of OSE cells to become deranged.

The underlying mechanism behind this hypothesis is that hormone oestrogen produced during ovulation is thought to promote this inclusion of OSE within the ovarian stroma at the formation of the described inclusion cysts and also affect mitotic activity of OSE causing unrepaired DNA damage and further leading to malignancy (Auersperg et al, 1998). To further support Fathalla's proposal it has been subsequently shown, that ovulation triggers inflammation responses and the release of several cytokines and chemokines and these may contribute to DNA damage of OSE cells which in that mitogenic environment may lead to neoplastic transformations (Ness et al, 1999). There is evidence which support that ovulation induction is implicated with ovarian carcinoma development whereas oral contraception and pregnancy as previously discussed has a protective role (Chene et al, 2009).

B) The gonadotropin stimulation hypothesis

This theory is based on the observation that the levels of the hormone gonadotropin are higher during the menstrual cycle when women ovulate than during pregnancy or while on oral contraceptive pill. Receptors for gonadotropin have been identified on OSE cells (Kang et al, 2001) and several studies suggest that synergistically with oestrogen, gonadotropin stimulates OSE proliferation and causes the induction of unrepaired DNA damage leading to malignancies (Konishi et al, 1999). Gonadotropin stimulation of human OSE cells can lead to increased proliferation and DNA synthesis as well as decreased cell death (Edmondson et al, 2006). The latter study also reported different responses from OSE derived from different women and suggested that the risk of EOC development may be dependent on the ability of OSE cells to respond to elevated levels of gonadotropins (Edmondson et al, 2006). The gonadotropin theory is also supported by the observation of increased ovarian cancer risk in postmenopausal women and women with polycystic ovaries where in both cases the levels of gonadotropin are especially high (Rossing et al, 1994)

C) The hormone stimulation hypothesis

Several studies have proposed the influence of other hormones to the development of epithelial ovarian cancer linked with OSE. Androgen levels when elevated have been linked with an increased risk with EOC as the OSE has an enzyme that can convert a weak androgen hormone, androstenedione to testosterone. This conversion leads to elevated levels of androgen in the plasma during ovulation but also at post-menopausal women and women with polycystic ovary syndrome. The OSE within inclusion cysts also gets over exposed to paracrine androgen hormones rather than circulating androgens and the increased incidence of EOC when these androgen levels are elevated suggests a role for the OSE in EOC development (Risch, 1998). OSE also has been reported to exhibit stimulation of DNA synthesis and proliferation after exposure to androgens *in vitro* (Edmondson et al, 2002).

Progesterone is a hormone that has been demonstrated to have a protective effect for EOC. This was based on the observation that pregnant women although having elevated levels of androgens in their plasma, they have even more elevated oestrogen and progesterone levels. Following the discovery of progesterone receptors on the ovarian epithelium, defects of the expression of these receptors has been suggested to lead to increased ovarian carcinogenesis (Risch, 1998). Another study suggested that progesterone prevents increased proliferation of the OSE by showing that progesterone receptor antagonists caused increased proliferation of OSE (Ivarsson et al, 2001).

1.3.2.1 Molecular and phenotypic evidence supporting OSE origin

The presence of cortical inclusion cysts lined with ovarian surface epithelium in patients diagnosed with epithelial ovarian cancer suggests that these germinal inclusion cysts may play a role in ovarian carcinogenesis (Mittal et al, 1993). Epithelial inclusion cysts of *BRCA1* and *BRCA2* mutation carriers display frequent accumulation of p53 which is known as the neoplastic p53 expression signature (Kerner et al, 2005, Pothuri et al, 2010) and another earlier study has shown that p53 expression was significantly elevated in inclusion cysts from patients with serous adenocarcinoma compared to normal and borderline adenocarcinoma (Hutson et al, 1995). This unique expression signature in the ovarian epithelial inclusion cysts is a representation of the oncogenic stress that causes increased growth signals and DNA damage which in turn lead to accumulation of p53. The frequent mutation of *TP53* seen in high grade serous carcinomas may be explained if their precursor lesions are the epithelial inclusion cysts.

The relationship between the expression of homeobox genes implicated in Mullerian development and lineage differentiation in EOC has also been demonstrated in studies where OSE cells were transfected with *HOXA9*, *HOXA10* and *HOXA11* and gave rise to serous, endometrioid and mucinous tumours respectively, suggesting that alterations in the HOX gene family in OSE cells may be an early event in tumourigenesis (Cheng et al, 2005, reviewed in

Naora 2007). The expression of the HOX genes in OSE that shares a common origin with the Mullerian derived tissues in development may be assigning a progenitor cell role to uncommitted OSE into adapting a Mullerian morphology and that might be answering to the paradox that EOC subtypes resemble more the Mullerian epithelial tissues.

1.3.2.2 The role of epithelial-to-mesenchymal transition and mesenchymal-to-epithelial transition in OSE derived carcinogenesis

Epithelial to mesenchymal transition (EMT) is a fundamental and conserved process during embryonic development, remodelling and restitution of tissue and wound repair. During these processes loss of cell-cell adhesion and increased cell movement is achieved and is characterised by distinct morphology of epithelial cells that acquire a mesenchymal phenotype (Thieri et al, 2003). The observed loss of differentiation, increased cell motility and invasion during EMT are characteristics of malignancy and it is now well established that carcinoma cells (epithelial derived tumours) are exploiting this mechanism to assimilate an increasingly aggressive phenotype especially relating to invasion and metastasis (Arias el al, 2001).

The molecule that linked EMT with carcinoma formation is E-Cadherin, a cell adhesion marker that is maintaining the adherent junctions of cells and guarantees their polarization and architecture of epithelial tissue. Loss of E-Cadherin is defining the development of epithelial tumours by increasing the invasiveness of cells and decreasing the intercellular adhesiveness. The important role that EMT has in carcinoma progression is supported by reports of several transcriptional repressors of E-cadherin such as Snail and Slug playing a key role in invasion and metastasis (Perl et al, 1998, Yang et al, 2005). The process of EMT is defined by E-Cadherin reduction and it’s been demonstrated to contribute in carcinoma development for colorectal (reviewed in Bates, 2005), gastric and breast cancer (Yang et al, 2005).The cells of origin for most epithelial cancers are more differentiated than the resulting tumours.

This is where the ovarian cancer paradox lies, with EOCs exhibiting an architectural organisation and appearing more differentiated than the proposed

OSE precursor cells. This presented the biggest argument regarding the origin of ovarian cancers along with the diversity of the subtypes resembling other reproductive tract organs. OSE, unlike other epithelial cells, have an uncommitted phenotype with mixed epithelial (keratin expression) and mesenchymal morphology (MUC-1, Vimentin, smooth muscle actin α) that differentiates into different types in response to environmental signals (Auersperg et al, 2001). Additionally, they rarely express E-Cadherin. During post-ovulatory remodelling of the ovarian surface, OSE cells migrate from the ovarian surface to the cortical stroma or are trapped in the ruptured follicle and undergo EMT under the influence of factors like EGF (Epidermal Growth Factor) and TGF-β (Transforming Growth Factor), IL-6 and other cytokines which are all well characterised EMT inducing factors (Salamanca et al, 2004, Zavadil and Böttinger, 2005, Ahmed et al, 2006).

The mesenchymal migratory characteristics conveyed by EMT to OSE cells at the position of the wound repair are for the purpose of repairing the scarred OSE and for allowing the OSE cells migrated within the stroma to incorporate with the stromal cells. The OSE cells that fail to undergo EMT are forming 'foreign' epithelial cell aggregates within the stroma resulting in the formation of the inclusion cysts which have been deemed as precursor lesions of ovarian carcinomas as discussed before. It is conceivable that the OSE lining the inclusion cysts after prolonged contact with all the EMT inducing factors within the ovarian stroma, undergo EMT, which similarly to other epithelial cancers may lead to neoplastic transformation.

It is interesting that the markers CA125 and E-Cadherin are expressed by the OSE cells lining of the inclusion cysts but not from OSE and both E-Cadherin and CA125 are markers found in EOC samples (Auersperg et al, 2001). The expression of E- Cadherin in tumour progression is controversial as there are reports that show that early (benign and borderline) ovarian tumours express E-Cadherin but the expression decreases as the tumours become more advanced (malignant) (Sundfeldt, 2003, Yoshida et al, 2009) but there is also evidence suggesting that the expression of E-Cadherin during cancer progression may be dynamic and highly contextual (Kowalski et al, 2003). This observation has set a role of mesenchymal to epithelial transition in the

progression of EOCs and could suggest an explanation to the paradox of the well-differentiated EOCs compared to other epithelial cancers. To summarise the proposed model, uncommitted OSE cells do not express E-Cadherin and if they fail to undergo EMT post ovulation they form inclusion cysts during which formation they undergo mesenchymal to epithelial transition (MET) and express E-Cadherin and also CA125. The epithelial lining of the ICs exposed to stromal stimulators is in turn undergoing EMT leading to tumour progression and advanced or secondary tumours can switch back to MET a process seen in other cancer metastasis as well (Okamoto et al, 2009). CA125 is also a marker that is expressed in the epithelium of the Mullerian ducts during development and this Mullerian type differentiation-metaplasia that OSE undergo in an adult female can help in understanding the development of so many different EOC subtypes. This may well be the indication that the Mullerian metaplasia to malignant progression is regulated by the microenvironment of the tumour and can exploit EMT or MET accommodated by the plasticity of OSE cells to easily differentiate.

1.3.3 Fallopian tube and other cell type origins for EOC

The mucosal epithelium of the fallopian tube (FT) consists of two cell types: ciliated and secretory. The ciliated cells are responsible for transportation of the ovum, sperm cell and the zygote. Ciliated cells express LhS28, a marker of basal bodies within the cilia, and Fox1, a transcription factor involved in ciliogenesis. The secretory cells secrete mucous that delays the progression of the spermatozoa through the fallopian tube. The secretory cells express Bcl-2 a mitochondrial suppressor of apoptosis and PAX8 a transcription factor involved in the development of reproductive tract (Levanon et al, 2010). Early serous carcinogenesis in the fallopian tube has presented elevated levels of another transcription factor seen in reproductive tract development, PAX2 (Chen et al, 2010).

Previous studies have tried to isolate epithelial cells of the fallopian mucosa using methods such as enzymatic treatment with trypsin and/or pancreatin. The epithelial status of the isolated cells was checked with cytokeratin antibodies of CK7, CK8, CK18, CK19 and PKK1. The lifespan of the

isolated cells was reported to be extremely short and those studies have failed to successfully subculture FT isolated cells (Henriksen et al, 1990, Comer et al, 1998). In a recent study Levanon et al have created the first reported FT model for studying EOC development where ex vivo cultures of secretory and ciliated epithelial cells of the fallopian mucosa were created to recapitulate FT function *in vivo* (Levanon et al, 2010). No study to date has successfully subcultured and fully characterised an *in vitro* FT model.

1.3.3.1 Molecular and phenotypic evidence supporting FT origin

Lately there is increasing debate about whether the fallopian tube is the predominant site of the origin for HGSCs. This debate initially derived from the morphologic discontinuum of the OSE with high grade SCs which resemble more the fallopian tube epithelium. This observation together with data that fail to assign a precise percentage of serous carcinomas arising from the OSE (Deligdisch et al, 1999) has led to more thorough investigation for the origin of HGSCs.

In the quest for a link between early malignancies in the FT and EOC development several studies have reported arguments to support this theory. The expression profiles of high grade SCs and the FT epithelium from *BRCA1* mutation carriers have been compared and exhibit striking similarities but differ to the expression profile of ovarian surface epithelium (Tone et al, 2008). Several studies have shown that in *BRCA1* mutation carriers who have undergone prophylactic surgery several were found to have an early malignancy, termed as tubal intraepithelial carcinoma (TIC) as this malignancy in 100% of the cases was found to originate in the fimbrial/distal region of the fallopian tube (FT) (Piek et al, 2003, Medeiros et al, 2006, Callahan et al, 2007). Other studies have shown that ~70% of early malignancies in women that underwent prophylactic surgery were neoplastic transformation of the salpingeal mucosa (Powell et al, 2005, Kindelberger et al, 2007). Mutational analysis in TIC revealed *TP53* mutations further supporting the distal fallopian tube origin for EOC (Kindelberger et al, 2007). Several studies have also shown a dominant tubal p53 signature of about 40% of cases from prophylactic surgeries (Folkins et al, 2008), with 95% located at the distal/fimbrial end of the fallopian

tubes (Leonhardt et al, 2011) and more specifically within segments of secretory cells (Lee et al, 2007).

Kindelberger et al have taken recent research into account to draw a preliminary model of pelvic carcinogenesis. In Figure 1.4 the proposed models of pelvic carcinogenesis including ovarian cancer are shown. One pathway shows the traditional Mullerian origin where OSE contribute in the formation of Mullerian inclusions in the ovary leading into carcinogenesis. The second scenario proposes that early malignancies are formed in the fallopian tube and these clones later invade the ovarian surfaces leading into carcinogenesis. Additionally within that scenario, following endometriosis the endometrial cells are transported ectopically to the ovary to give rise to endometrioid carcinomas (Kindelberger et al, 2007, reviewed by Dubeau, 2008).

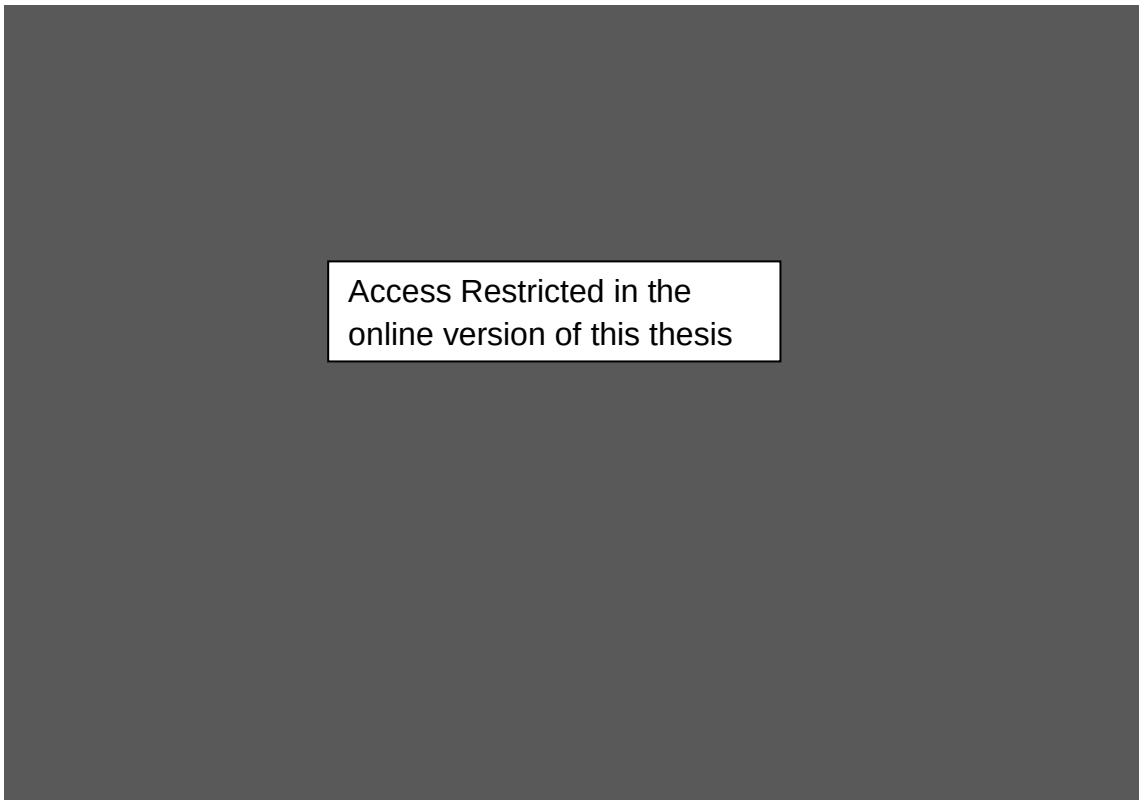


Figure 1.4**:** Model proposing a multiplicity of sites of origin for the pathogenesis of EOC. Some tumours may arise from the OSE through Mullerian metaplasia, some after ectopic transportation of tumour cells from TIC in the distal end of the fallopian tubes and others from another source close to the ovarian surface such as the endometrium (Kindelberger et al, 2007).

1.4 Animal and *in vitro* models used to study EOC

Experimental animal models for human disease and cancer are essential to help elucidate the biological and genetic factors that govern the phenotypic characteristics of the disease. Animal models have the advantage of studying the disease in the mammalian environment where the mechanisms of the effects of steroids and hormones can be better studied in regards to EOC initiation.

There are three types of animal models for EOC tumourigenesis (reviewed in Vanderhyden et al, 2003): i) Animal models that develop ovarian tumours spontaneously such as hens and some strains of mice, but their low incidence make them a poor tool for experimental studies. ii) Animal models with xenografts of OSE cells transformed *in vitro*. Although this approach has allowed for evaluation of oncogenes that contribute to the development of EOC, it can also be used in the investigation of the initiation and early events of ovarian tumourigenesis. iii) The development of transgenic animals that will develop EOC with approaches to transform their ovarian surface epithelium, to create defined genetic lesions that could be studied at various stages as they inevitably develop ovarian cancer *in situ*.

The first *in vitro* tool for studying molecular markers thought to be implicated with ovarian cancer development and progression as well as chemoresistance, are ovarian cancer cell lines that have been established by primary ovarian tumours or ascites. However, a limitation of this approach, though cheap and easily controlled, is that there is always the risk that alterations observed might have been introduced through prolonged tissue culture and cannot resemble the microenvironment in the tumour.

Another *in vitro* model for studying EOC transformation is achieved by introducing the molecular alterations to study in primary normal ovarian surface epithelial (NOSE) cell lines from the OSE as it is one of the proposed cells of origin for EOC, known as transformation models. Using primary cell lines is challenging as they have a limited lifespan in culture using hTERT (Yang et al, 2007). Attempts to create *in vitro* transformation models using immortalised NOSE cells with knocked down *TP53* and *Rb* have been made but did not lead into a neoplastic phenotype (Yang, 2007).

However, all these studies have used *in vitro* two dimensional (2D) cell models. 2D cell cultures are cell monolayers that do not exhibit the complex cell - cell and cell - extracellular matrix (ECM) interactions that are observed in tissue *in vivo*. This is a major limitation, so 3D culturing of cells can be used to try and re-capitulate tissue architecture and thus improve the biological relevance of *in vitro* models. There are several studies that have provided substantial evidence proposing that 3D cultures more closely resemble the *in vivo* architectural organisation and microenvironment than 2D cultures, and that culturing cells in 3D can lead to phenotypic and molecular changes that reflect the *in vivo* biology of the cell more closely than 2D cultures (Knuechel et al, 1990, Ghosh et al, 2005, Zietarska et al, 2007).

In an attempt to create an *in vitro* model to resemble the genetic heterogeneity in EOC and study the initiation and early stage neoplastic progression an *in vitro* transformation model has been created within our group by firstly creating three dimensional (3D) models of normal OSE (NOSE) cells, which better resemble the microenvironment of the ovary. Lawrenson et al subsequently created 3D models of partially transformed immortalised NOSE (IOSE) cell lines by overexpressing the *MYC* oncogene to later create double mutant cell lines by subsequent expression of *KRAS* and *BRAF* and showed neoplastic progression *in vitro* and differential gene expression profiles between the transformed cell lines (Lawrenson et al, 2009, Lawrenson et al, 2011).

1.5 Genetics of epithelial ovarian cancer

1.5.1 Heritability in Ovarian Cancer

EOC is a polygenic disease that can be caused by mutations, alterations and duplications of genes and 90% of the cases are sporadic with no family history. The remaining 10% of cases are inherited and this is termed as familial ovarian cancer (Holschneider and Berek, 2000).

Family studies have well-established that the cumulative risk of developing ovarian cancer for a woman with a family history of ovarian cancer is significantly higher than women of the general population where the risk is estimated as 1% by the age of 70. An extensive meta-analysis of such studies has been performed by Stratton and colleagues, which reported that the relative risk of a woman with one affected first degree relative is 3.1% (95% CI 2.6-3.7), with the risk being higher to sisters and daughters, 3.8 and 6 respectively, than to mothers 1.1%, the nature of the difference is still not understood. The risk increases for women with two or more affected first degree relatives to 12% but with very wide 95% CI (Stratton et al, 1998).

Even though family studies provide information about the familial incidence of ovarian cancer, they fail to establish whether the observed increased risk is due to genetic factors. Twin studies, however, provide valuable information on whether familial risk is due to shared genetic factors rather than shared environmental factors. Twin studies, compare monozygotic twins that arise from a single fertilised egg and have 100% identical genotypes and dizygotic twins that share only 50% of their genotypes. Both groups share the shame environmental similarities and comparing the similarities of monozygotic (MT) with dizygotic twins (DT) can provide valuable observations on the influence of genetics in inherited risk of cancer. The principle behind this is that in order to establish whether familial aggregations of cancer are hereditary or due to shared environmental factors one needs to examine whether the observed cancer phenotype is concordant with the observed genetic similarities and in that case MT will more frequently have a cancer in the same anatomical site than DT (Neale et al, 1992).

An extensive comparison of MT and DT cancer concordance in large twin cohorts has been performed by Lichtenstein et al in 2000. This study reported a higher concordance of the same cancer in MT than DT in stomach, colorectal, lung, breast, prostate and ovarian cancer indicating that genetic factors have a more important role in cancer predisposition than non genetic factors shared within families in a dominant genetic effect manner. The cumulative risk for EOC development was calculated as 22% which is 6 fold higher than the risk shared between non-twin siblings (Lichtenstein et al, 2000).

Following the realisation that familial aggregates of cancer are characterised by similar genetic contributions and after the generation of the human genome reference sequence enabled to distinguish genetic variation in distinct populations (The International HapMap, 2005) many probable genetic contributions to ovarian cancer susceptibility were discovered. Familial ovarian cancer has been categorised into two distinct familial cancer syndromes which are inherited in an autosomal dominant manner, hereditary breast/ovarian cancer syndrome and Lynch syndrome or hereditary non-polyposis colorectal cancer.

1.5.1.1 High susceptibility/penetrance genes for ovarian cancer

In hereditary cancer highly penetrant tumour suppressor genes like *BRCA1/2, MSH* are dominantly 'acting' as the inheritance of only one faulty gene copy is sufficient to predispose an individual to cancer, thus these genes are referred to as dominantly inherited. However, they 'act' recessively at the cellular level, so mutation or alteration of the normal copy is required for cancer to occur. This is better explained by the proposed genetic model of cancer development by the inactivation of tumour suppressor genes by Knudson et al in 1971 which is known as the two hit hypothesis. The first hit is the passing of a mutated copy of a gene to the offspring and the second hit is a random somatic mutation or deletion or loss of heterozygosity leading to loss of the normal second copy of the gene (Huang et al, 1997). In sporadic cancer cases, which will be later discussed, mutation in both copies of the gene have to randomly occur, thus families carrying an already defective copy of the gene have a

higher predisposition as only one random event needs to occur for the gene to be completely defective (Figure 1.9, A)

1.5.1.2 Breast/Ovarian Cancer Syndrome associated genes

The fact that ovarian cancer has a high prevalence in certain families led to the investigation of inheritance of predisposing mutations in cancer susceptibility genes within the familial ovarian cancer cases and the first tool used to link a gene to the disease was genetic linkage studies.

Genetic linkage can be determined in humans using pedigree analysis. The first convincing evidence of a breast/ovarian cancer associated gene and its localisation was reported in the early 1990s with 40% of the families tested linked to a marker in chromosome 17q (Hall et al, 1990, Narod et al, 1991). These initial observations were confirmed by a larger study on 214 families, which provided evidence for linkage of breast and ovarian cancer of the majority of the families to a marker in the same chromosomal location (Easton et al, 1993). The *BRCA1* gene was later identified and described by positional cloning. It's exact location is on chromosome 17q12-21 and is composed of 22 coding exons spanning over 100 Kb of genomic DNA and coding for a protein of 1863 aa and 208 kDa in size. Germline mutations of *BRCA1* were detected in five out of eight families (Miki et al, 1994). A year after the discovery of *BRCA1* as a breast/ovarian cancer high susceptibility gene, *BRCA2* was localised in chromosome 13q12-13 and was linked with 74% of the families studied (Wooster et al,1994).

Several studies have since reported germline mutations in the *BRCA1* and *BRCA2* genes with prevalence estimates varying considerably. Reports vary for *BRCA1* from as high as 82% (Narod et al, 1995), to 57% (Ford et al, 1995) and 43% (Gayther et al, 1999). The most comprehensive analysis of *BRCA1* and *BRCA2* mutations in ovarian cancer to date evaluated 283 families with 2 or more first degree relatives with ovarian cancer and reported a 37% prevalence for *BRCA1* and 9% prevalence for *BRCA2* (Ramus et al, 2007). Interestingly, prevalence of *BRCA1* and *BRCA2* mutations is strongly associated with the extent of ovarian cancer family history with the risk

increasing in multicase families as reviewed in (Gayther and Pharoah, 2010, Pharoah and Ponder, 2002, Ramus and Gayther, 2009).

The variance in the estimates for penetrance for *BRCA1* and *BRCA2* in familial studies can be attributed to an extent to allelic heterogeneity but the magnitude of the differences reported is also suggesting that other factors could be modifying the risk between carriers. These could be lifestyle and environmental factors such as low parity increasing the risk, and oral contraceptive use reducing the risk. Additionally, genetic alterations in genes such as the androgen receptor gene and *HRASI* proto-oncogene have been shown to have an effect in modifying the risk for ovarian cancer (reviewed in Pharoah and Ponder 2002).

A meta-analysis of 10 studies reported a cumulative risk for ovarian cancer to the age of 70 years of 40% and 18% for *BRCA1* and *BRCA2* respectively (Chen et al, 2006). Another study combined data from 22 different studies and has reported the cumulative risk for *BRCA1* mutation carriers from population based studies to be 39% and 11% for *BRCA2* mutation carriers (Antoniou et al, 2003).

The spectrum of *BRCA1* and *BRCA2* mutations has been reviewed by Pharoah and Ponder in 2002. More than 250 different germline *BRCA1* mutations had been discovered with the most frequents being 185delAG and 5382insC both founder mutations in the Ashkenazi Jewish population. The most frequent of the germline alteration for *BRCA2* is 6174delT, common in Ashkenazi Jews, and 999del5. Most germline *BRCA1* and *BRCA2* mutations are predicted to result in protein truncation caused by frameshift, nonsense, splice-site and regulatory mutations and large genomic deletions, all distributed along the length of the genes. In-frame deletions and missense mutations are more uncommon in *BRCA1* but found in *BRCA2*.

The association of ovarian cancer risk with the *BRCA* genes has made critical the investigation of their function. In the last years significant progress has been made regarding the elucidation of BRCA1 and BRCA2 function at the cellular level. It is now widely established that *BRCA* genes act as tumour suppressor genes and have a role in maintaining genomic stability/integrity. Profound chromosomal translocations, duplications and aberrant fusion events

between non-homologous chromosomes have been observed in *BRCA1/2* deficient cells (Moynahan et al, Feb 2001, Moynahan et al, June 2001). Findings that BRCA1 and BRCA2 nuclear foci formed after treatment with DNA-damaging agents interact with RAD51, an essential protein for the repair of double-strand DNA breaks, are suggesting that both genes are important for genomic integrity through a direct role in DNA repair (Chen et al, 1998, Scully et al, 2007).

The clinical features and outcome of ovarian cancer patients with and without *BRCA1/2* mutations has been investigated but results of such studies have been controversial. Significantly prolonged survival for patients and higher proportion of serous adenocarcinomas was reported for *BRCA1* mutation carriers compared to non-carriers in some early studies (Rubin et al, 1996, Aida et al, 1998). These studies were methodologically criticised for bias (Cannistra, 1997 (letter)) but were supported with subsequent studies (Boyd et al, 2000, Ben et al, 2002). A Swedish study has reported higher proportion of serous adenocarcinoma in *BRCA1* mutation carriers but found their survival similar or worse compared to non-carriers (Johansson et al, 1998). Pharoah et al reported no survival difference in mutation carrier families and Ramus et al reported slightly improved survival in Ashkenazi ovarian cancer patients with founder *BRCA1* and *BRCA2* mutations but not statistically significant (Pharoah et al, 1999, Ramus et al, 2001). A recent larger and conclusive study has shown that *BRCA1* and *BRCA2* mutation carriers have improved prognosis with *BRCA2* mutation carriers showing the most improved prognosis (Bolton et al, 2012).

The mechanism that underlies the reported improved survival for *BRCA* mutation carriers could potentially be a better response to the chemotherapeutic drugs that act by inducing DNA damage. The proposed role of *BRCA* genes in DNA repair mechanisms could explain why their loss is triggering a more efficient death of the treated cells due to the induced DNA damage, but this will be further discussed later.

1.5.1.3 Lynch Syndrome or HNPCC associated genes

Hereditary non-polyposis colorectal cancer (HNPCC) is caused by germline mutations in genes involved in the DNA mismatch repair pathway, including *MLH1, MSH2, MSH6* and *PMS2.* HNPCC is associated with tumours that exhibit microsatellite instability (MSI). The genes of the mismatch repair pathway (MMR) have the role of spotting and repairing any mismatches of single nucleotides as well as insertion or deletion loops. Defects in those genes can lead to accumulation of somatic mutations in a cell and cause tumourigenesis. Microsatellite markers are sequences of DNA containing repetitive nucleotides and they are very susceptible in gaining errors if there is a defect in the MMR mechanism. Thus, in Lynch Syndrome where MMR genes are defective the number of microsatellite repeats observed in tumour is not consistent with normal tissue, and this condition is called microsatellite instability (MSI) (Gruber and Kohlmann, 2003, Manceau et al, 2011).

The life-time risk for ovarian cancer development is 12% within families with Lynch syndrome (Aarnio et al, 1999). The majority of Lynch syndrome is attributed to mutations of *MLH1* and *MSH2* genes with 50% and 40% prevalence respectively. MSH6, MSH3, PMS1 and PMS2 mutations account for the rest 10% of the cases (Wheeler et al, 2000). The cumulative risk for developing ovarian cancer for Lynch syndrome mutation carriers has been estimated with the genotype restricted likelihood, which accounts for ascertainment bias in 537 families. The cumulative risk by the age of 70 for MLH1 mutation carriers was reported to be 20%, for MSH2 24% and for MSH6 1% (Bonadona et al, 2011).

1.5.2 Moderate/Low susceptibility genes for ovarian cancer

With less than 60% of the familial ovarian cancer risk due to mutations in the high penetrance genes *BRCA1*, *BRCA2* and MMR genes, this provided evidence of the existence of other genes which confer moderate or low risk in the development of ovarian cancer, indicating that ovarian cancer is a polygenic disease influenced by many genes. The remaining susceptibility genes could be of moderate or low penetrance with a combined risk of 50% and could

account for the remaining excess familial risk and also attribute the risks for sporadic ovarian cancer in the population. This polygenic model of ovarian cancer is supported by evidence that was reported by the twin studies which showed that genetic factors are more important than environmental factors shared by the twins in ovarian cancer development. The tool used by geneticists to identify moderate/low risk susceptibility alleles is genetic association studies.

1.5.2.1 Genetic association studies

The aim of genetic association studies is to find associations between genetic polymorphisms and a disease or other quantitative characteristics by studying unrelated subjects in search of common variants that would elucidate cancer incidence in the general population. We have seen that linkage studies were used to identify the *BRCA1* and *BRCA2* genes in ovarian cancer families. Association studies are fundamentally different from linkage studies in the sense that in association studies the same allele is associated with the disease across the whole population whereas linkage studies present the association of different alleles with the disease across different families. Since the realisation that ovarian cancer is a complex disease that involves many genes with a small effect, termed as low/moderate susceptibility genes, the shift of genetic epidemiology to association studies was evident. Association studies have a greater power to detect small effects than linkage studies do, but require many markers/polymorphisms to be analysed in large number of presumably unrelated individuals (Risch et al, 1996). The importance of genetic association studies in genetic epidemiology is explained by the need to identify more low/moderate penetrance alleles, the identification of large numbers of single nucleotide polymorphisms (SNPs) within the human genome, and the advances in genotyping techniques .

Another form of association studies are survival association studies where genetic variants are screened and compared between alive and deceased patients in order to identify whether they may have an effect on the patients' survival time and also the response to the treatment, which can lead into more personalised treatment.

1.5.2.2 Single nucleotide Polymorphism (SNP)

Genetic association studies try to identify an association between polymorphisms and susceptibility in cancer or other diseases, by typing single nucleotide polymorphisms (SNP) in a population. A SNP represents DNA sequence variation in a single nucleotide with a >1% population frequency. Across the human genome the number of SNPs represents the majority of the polymorphisms. The International SNP Map working Group had provided a map of 1.4 million candidate SNPs and observed that two haploid genomes have a difference of 1 nucleotide per 1.331 bp, which is turn infers to the existence of 11 million SNPs in a genome of 3.2 billion bp that vary in at least 1% of the entire population (Wang et al, 1998, Kruglyak and Nickerson, 2001, The International SNP Map Working Group, 2001). However, with the release of the 1000 genomes project the number of mapped SNPs reached around 15 million (The 1000 genomes project consortium, 2011). SNPs are predominantly biallelic, meaning they are two different copies of a SNP in the human chromosome. One of the alleles of a SNP is maternally and one paternally derived as seen at Figure 1.5.

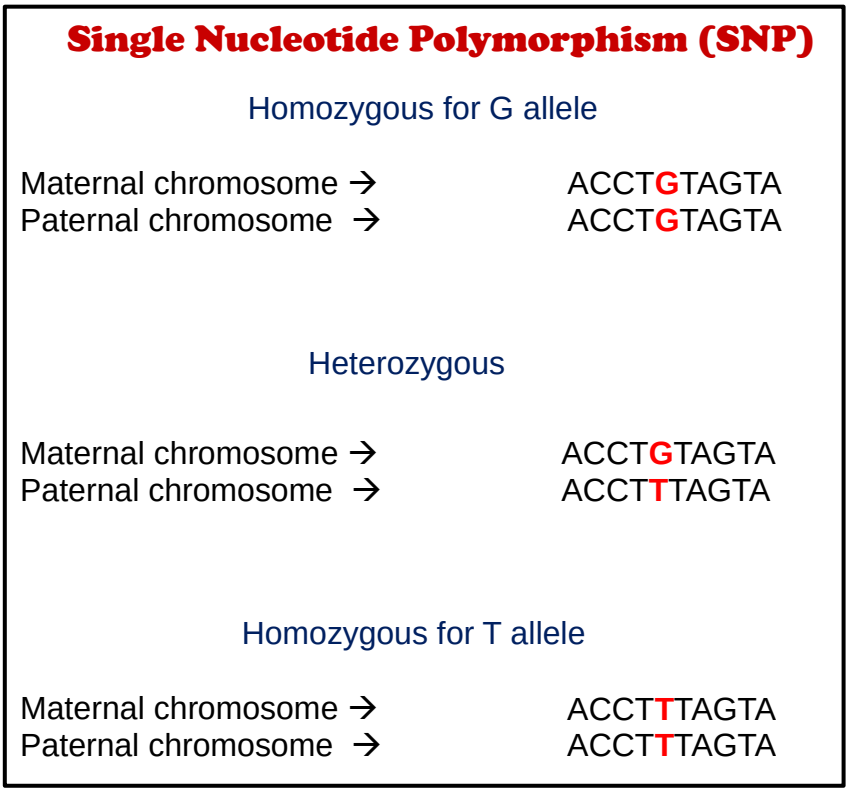


Figure 1.5: Possible variations of a SNP in population.

SNPs can have a functional role in influencing human phenotype and can be divided into two categories; the coding and noncoding SNPs. Coding SNPs are located in an exon of a gene and are termed synonymous if they do not modify the polypeptide sequence of the gene product due to the redundancy of the genetic code; for example a **TC** SNP can result to codons TT**T** or TT**C** and both code for the amino acid phenylalanine. Functional roles by synonymous SNPs have been regularly reported, with some examples being reports of synonymous SNPs having a role in changing the specificity of the resulted protein of the *Multidrug Resistance 1 (MDR1*) gene (Kimchi-Safati et al, 2007) and affecting gene expression and mRNA folding and processing (Wang and Sadée, 2006). Synonymous variations have also been associated with reducing the amount of the resulting protein by altering the stem loop secondary structures of the mRNA of *Catechol-O-methyltransferase* gene (Nackley et al, 2006) and numerous other reports have assigned similar functions in altering the availability of the protein by changing the structure of the protein or by alternate splicing due to synonymous SNPs. A recent study has reported the implication of synonymous SNPs in altering binding capacity of (micro RNAs) miRNAs in a way that the protective allele was down regulated by this miRNA and the risk allele was not allowing the binding of the miRNA (Brest et al, 2011).

Coding SNPs are termed non-synonymous when they do change the amino acid sequence of the protein and thus having the obvious functional role in protein production. A non-synonymous SNP can either be missense, which results in an alternative amino-acid, or nonsense which results in a production of a truncated protein by introducing a premature stop codon. About 60000 SNPs exists in the coding regions of approximately 30,000 genes (The International SNP Map Working Group).

Non-coding SNPs may be located in the gene's introns, untranslated regions, or other regulatory regions but may be functional in influencing the availability of a protein by altering sites for transcription factors or disrupting splice sites on the translational level (Reumers et al, 2008). SNPs can also be found inbetween genes and in gene deserts which are large genomic segments with no genes and limited sequence conservation. However, gene deserts

rarely have conserved regions which could harbour transcriptional regulatory elements and influenced by the non- coding SNPs located there.

When a disease-associated SNP is identified it is important to estimate the degree on which the amino-acid substitution is damaging in cases of non-synonymous SNPs and also the predicted function of synonymous or non-coding SNPs if based in a regulatory region of the gene. There is a lot of progress recently in building functional prediction databases for SNPs. Databases, such as PupaSuite, can be used to evaluate known or predicted functional information about all the categories of SNPs (Barrett et al, 2005). Non-synonymous SNP functional prediction databases also include SIFT, PMUT, SNPeffect, LS-SNP, SNAP, PolyPhen, TopoSNP (reviewed by NG. and Henikoff, 2006). A recently published database for non-synonymous SNPs' functional predictions is providing the most comprehensive integrated database of functional predictions from multiple algorithms (Liu et al, 2011).

1.5.2.3 Linkage disequilibrium

A genotyped SNP might be either the putative causal variant (direct association) or can be correlated with the SNP that is the actual causal variant of the studied phenotype (indirect association). It is now well established that SNPs within a chromosomal locus can be in correlation with each other. This non-random correlation between SNPs is called linkage disequilibrium (LD). Pairwise LD between two SNPs is measured by Lewontin's D, known as association probability, but more commonly LD is expressed by the index r^2 which is related to the allele frequencies and is the square of the conventional correlation coefficient between the allele at the typed locus and the allele at the causal locus. LD is also relevant to haplotype blocks (haplotype is the allelic configuration along a single chromosome), by characterising the haplotype diversity within these blocks (Cordell and Clayton, 2005). The process of tagging SNPs relies on LD since the causal variant does not need to be studied directly but information can be provided by a SNP that is in LD with it. This will be discussed in more detail later in the description of genetic association studies designs.

1.5.2.4 The International HapMap project and the 1000 genomes project

Of the >10million SNPs identified more than 5 million have been validated by HapMap. The International HapMap Consortium used DNA samples from Africa, Asia and Europe in order to investigate the common genetic variation in the human genome. The International HapMap project has genotyped millions of SNPs across the human genome and reported their frequencies in the population as well as the degree of LD between variants (The International HapMap, 2005).

More recently, the 1000 genomes project has provided a deeper characterisation of human sequence variation in the purpose of aiming association and functional studies. Recently published data have revealed the location, allele frequency and local haplotype structure of approximately 15 million SNPs. The vast majority of the variants were already present in dbSNP but a lot of newly discovered variants were identified mostly in relation to specific population groups with the African ancestry population providing the highest fraction of novel variants identified (The 1000 Genomes Project Consortium, 2011).

1.5.2.5 Candidate gene and pathway approach to identify low penetrance susceptibility genes for ovarian cancer

The first form of genetic association studies was to investigate the association of potentially functional SNPs, causing an amino acid change, with ovarian cancer development. In the design of these studies, SNPs within genes that have been predicted to have a functional role in the development of the disease for example from tumour expression studies are genotyped. Also SNPs within genes that are part of functional pathways that have been implicated in carcinogenesis, such as cell cycle control, DNA repair, steroid hormone metabolism, inflammation pathways. Direct association is then investigated for these SNPs with pathways by having them genotyped in a number of ovarian cancer cases and their genotypic frequencies are then compared with healthy individuals to identify possible direct association of the typed polymorphisms with the development of the disease.

Small number case-control studies (not exceeding 200 cases or controls) have reported weak ovarian cancer associations of SNPs within candidate genes or genes involved in candidate pathways with a role in ovarian cancer development. One SNP of the cell cycle regulator gene *TP53,* which is a prevalent gene alteration in ovarian tumours, was found to be associated with the risk of developing various neoplasias including ovarian cancer *(Agorastos et al, 2004).* Another association reported is between a SNP in *SOD2*, a gene involved in inflammation and oxidative stress which are processes associated with ovulation (Olson et al, 2004). Ovarian cancer is a hormone-associated cancer and genes involved in hormonal regulation have also been investigated. The functional variant A2 was typed in *CYP17*, a gene involved in steroid hormone biosynthesis and was associated with increased risk in ovarian cancer development (Goodman et al, 2001, Garner et al, 2002) whereas another study did not report such evidence (Spurdle et al, 2000). Additionally, strong association was found between two variants in the *FSHR* gene, which codes for a follicle stimulating hormone receptor in the ovaries (Yang et al, 2006). Functional SNPs within genes involved in homologous recombination of double strand DNA breaks have been genotyped in relatively larger number of samples (>1000 cases and >2000 controls) with the most relevant positive associations of SNPs N372H, R188H and rs1799796 within *BRCA2*, *XRCC2* and *XRCC3* respectively (Auranen et al, 2005). Most of the results of those studies have not been validated in larger sample sizes however.

Over time association with coding variants yielded only a handful of conclusive associations, mainly in bladder cancer, but not in other cancers including ovarian cancer. Very few coding variants were confirmed suggesting that the majority of common alleles contribute through alterations in regulation or expression of genes and pathways. So the putative causal SNP associated with risk in the development of a disease may possibly be a non-coding SNP and this realisation triggered the design of a second form of association studies where the polymorphism typed is a surrogate for the real causal SNP around the candidate locus and this is performed by using tagging SNPs (tSNPs).

In order to identify genetic variation without genotyping every SNP in a chromosomal loci, a representative SNP called the tagging SNP is used, which

provides genotype information about all the SNPs that are found in LD (r^2=1 is a marker of complete LD between 2 SNPs) with the selected SNP (Goldstein, 2001). The combination of the alleles in multiple loci is called the haplotype. HapMap provides extensive genotype data and this resource was used to select tSNPs. SNP tagging is then performed using the Tager component of Haploview that is a software that provides information about haplotype blocks and LD between SNPs (Barrett et al, 2005).

The indirect association approach using tagging SNPs was used by several studies that reported association of tSNPs with ovarian cancer risk in genes such as *SHMT1*, coding for an enzyme involved in one-carbon metabolism implicated with methylation patterns and DNA synthesis (Keleman et al, 2008), tumour suppressor gene *RB1* and genes in the mismatch repair pathway such as *PMS2* (Song et al, Nov 2006, Song et al, Oct 2006) and *MLH1* (Harley et al, 2008).

However, these studies produced only marginally significant results that could be false positives, even when the statistical power was strengthened by following consortium approaches. Consortium approaches are study designs within multicentre collaborations which perform genotyping of SNPs in a staged way and the positive associations that are found in early stages are then validated in a larger number of samples, which gives the results a greater statistical power and minimises the errors. There are several consortium approach studies that have reported ovarian cancer associations with putative causal SNPs within candidate genes such as cell cycle regulators *CDKN2A, CDKN1B* (Gayther et al, 2007), *TP53* (Schildkraut et al, 2009), tumour suppressor *RB1,* progesterone receptor *PGR* and cell cycle kinase *AURKA* (Ramus et al, 2008), *MSL1* which is a gene implicated in histone and chromatin activation (Peedicayil et al, 2010) and several others.

1.5.2.6 Ovarian Cancer Association Consortium

The initial reports of all association studies failed to establish strong association of markers with ovarian cancer risk. The main issues for the difficulty in conducting statistically and methodologically rigorous ovarian cancer association studies are the following:

i) Because of the rarity of ovarian cancer the number of subjects used in most studies is hundreds rather than the thousands needed.
ii) Because of the vast number of SNPs in the human genome the false positive associations using the significance level of 0.05.
iii) Epithelial ovarian cancer is a heterogeneous disease and the small number of cases in each histological group makes it difficult to stratify for subtype in order to establish whether a candidate polymorphism affect more the risk for development of one subtype or the other.
iv) Stratification of association data according to other epidemiological risk factors such as ovulation and endometriosis is difficult because of small number of cases.
v) Possible false positive associations in some studies arose from failure to perform stratification of population as the allele frequencies are different between ethnic groups and this may bias the results.

The demand for addressing all these challenges and for making progress into understanding the contribution of genetic polymorphisms to susceptibility in ovarian cancer has led to the establishment of an Ovarian Cancer Association Consortium (OCAC) in 2005. This collaboration brought together more than 20 groups around the world which together pooled more than 10,000 cases and 15,000 controls. The ultimate goal of the OCAC was to identify as many possible low/moderate susceptibility alleles to EOC and stratify ovarian cancer risk using the genetic variation associations in combination with known epidemiological risk factors such as family history, ovulation and parity, oral contraceptive use. Better screening and prevention will be the outcome of such a focused effort with the goal of minimising EOC incidence and mortality. The initial efforts of the OCAC was to validate initial associations between SNPs in candidate genes and ovarian cancer reported by individual groups using the indirect SNP tagging approach and analysing the results in order to be in line with the issues for a robust association study design mentioned above (Ramus et al, 2008).

1.5.2.7 Genome Wide Association Studies: An agnostic but powerful approach

Identifying a small number of SNPs that will together elucidate the risk for developing a disease in an estimated number of >15million SNPs in the human genome can be quite a challenge and taking into account that many of those SNPs may be non-coding, candidate pathway studies have a very limited power into achieving this goal. Genome Wide Association Studies (**GWAS**) have thus emerged as an important and powerful tool for evaluating the frequency of thousands of SNPs in the human genome without any selection based on function. This agnostic approach is using tSNPs (r^2>0.8) to discover regions that harbour genetic variants that are associated with the trait in question. It was predicted that a GWAS would require ~500,000 SNPs to be genotyped because of the low level of LD in the human genome (Kruglyak, 1999).

GWAS are now widely used to identify low/moderate risk susceptibility alleles in several types of diseases including cancer in order to establish associations of loci with the disease. The tSNPs are genotyped using high-throughput genotyping platforms in consortium collaborations. Some examples are loci associated with risk in breast (Hunter et al, 2007), colon (Tomlinson et al 2008), lung cancer (Amos et al, 2008, Hung et al, 2008) and prostate cancer (Thomas et al, 2008). An extremely significant result of GWAS was reported by the breast cancer-association consortium that identified an association of a SNP in *FGFR2*, a gene coding for fibroblast growth factor receptor, with increased risk for breast cancer with a P value of 2×10^{-76} (Easton et al, 2007)

GWAS are often performed using a staged study design where significant associations of each stage are taken to the next stage to be genotyped in larger numbers to test replication and ultimately identify SNPs that reach genome-wide levels of statistical significance (p value<10^{-7}). The results that are analysed in each stage are then combined in order to strengthen the power of the study and exclude false positives (Gayther and Pharoah, 2010). The ovarian cancer GWAS designed and carried out by our group and collaborators is shown in Figure 1.6.

The first results from this study reported an association of SNP rs3814113 in locus 9p22 (Song et al, 2009). Further analysis of the GWAS data

also stratifying by subtype identified five additional susceptibility loci at 8q24, 2q31, 3q25, 17q21 (Goode et al, 2010) and 19p13 (Bolton et al, 2010). EOC is a heterogeneous disease consisting of serous, endometrioid, clear cell and mucinous subtypes and the genetic alterations involved in each subtype are also distinct as discussed before. GWAS multicentre collaborations, due to the collective numbers of samples, allow the stratification of association analysis according to distinct subtypes. Such analysis revealed that EOC heterogeneity is also influenced by common low-penetrance genetic variation, with the strength of association stronger in serous-only ovarian cancer cases in five out of six loci, the difference being more profound in loci 8q24 and 19p13. Common genetic susceptibility loci identified by our group's and collaborators' ovarian cancer GWAS and their association with histological heterogeneity are presented in Table 1.1.

Locus	Associated SNP	SNP Location & Function	All Cases/ Serous only	Controls	Stages 1, 2 and 3 combined			
					All subtypes		Serous cases only	
					OR (95%CI)	P-value	OR (95%CI)	P-value
2q31**	rs2072590	Within non-coding gene, unknown	10,406/ 5,925	16,340	1.16 (1.12-1.21)	4.5×10^{-14}	1.2 (1.14-1.25)	3.8×10^{-14}
3q25 **	rs2665390	Gene intron, unknown function	10,406/ 5,896	17,369	1.19 (1.11-1.27)	3.2×10^{-7}	1.24 (1.15-1.34)	7.1×10^{-8}
8q24**	rs10088218	Intergenic non coding, unknown function	10,462/ 5917	16,362	0.84 (0.80-0.89)	3.2×10^{-9}	0.76 (0.70-0.81)	8×10^{-15}
9p22*	rs3814113	Intergenic non coding , unknown function	8,761/ 4,847	11,831	0.82 (0.79-0.86)	5.1×10^{-19}	0.77 (0.73-0.81)	4.1×10^{-21}
17q21**	rs9303542	Gene intron, unknown function	10,242/ 5814	13,091	1.11 (1.06-1.16)	1.4×10^{-6}	1.14 (1.09-1.20)	1.4×10^{-7}
19p13***	rs8170	Coding, non synonymous	13,115/ 5,556	13,115	1.12 (1.07-1.17)	2×10^{-6}	1.2 (1.13-1.27)	$6x10^{-10}$
	rs2363956	Coding, non synonymous	13,115/ 5,556	13,115	0.96 (0.89-0.94)	4×10^{-7}	0.96 (0.89-0.94)	7×10^{-11}

Table 1.1: SNPs found in association with epithelial ovarian cancer susceptibility. OR (95% CI): per allele odds ratio (95% confidence interval). The P values presented for all the samples and serous only samples are from the combined 3 stages of the GWAS *(Song et al, 2009), ** (Goode et al, 2010), *** (Bolton et al, 2010).

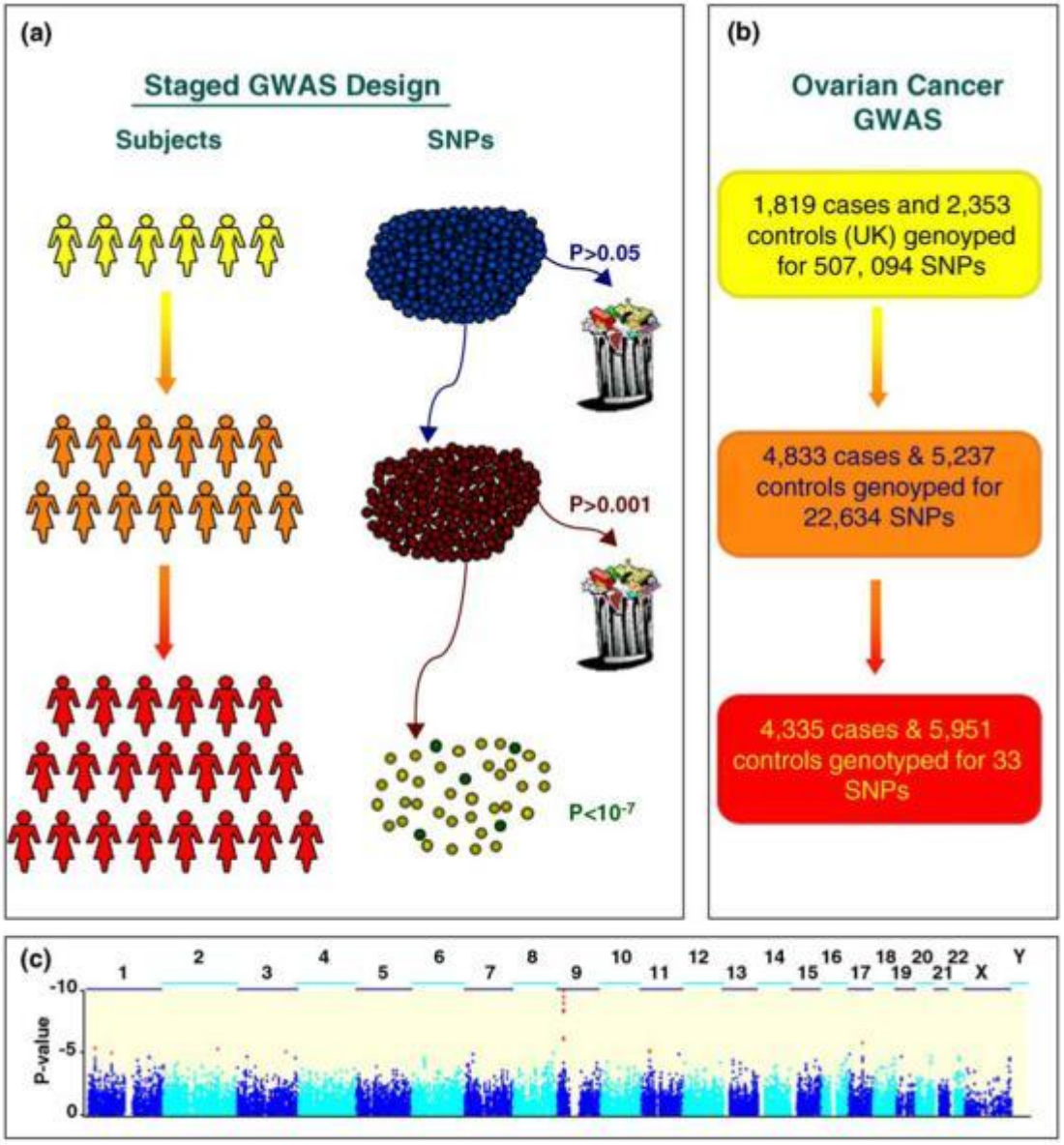


Figure 1.6: Schematic representation of GWAS design. a) The design of GWAS in the 3 stage approach. In the first stage a large number of SNPs is genotyped in a small number of samples, The SNPs that have a statistical significant value are then genotyped ion a larger number of samples/ And finally the few SNPs that have a P>0.001 after combined analysis of stage 1 and 2 are taken forward to be genotyped in the whole sample set. b) The ovarian cancer GWAS designed by the Ovarian Cancer Association Consortium is shown. C) A Manhattan plot is showing the combined stage 1 and 2 significant SNPs. Red highlights the SNPs found to have significant association with ovarian cancer with P values<10^{-5}. (Figure obtained from Gayther and Pharaoh, 2010, *Permission to reproduce this image has been granted by Elsevier using Copyright Clearance Center's RightsLink service)*.

1.6 Post GWAS characterisation: Functional follow up of susceptibility loci

Functional studies are often necessary to elucidate the biological mechanisms by which an associated region confers disease susceptibility. The challenge in order to translate the implication of a susceptibility locus found to be associated with risk in EOC is to identify how the candidate associated SNPs affect risk modification and also identify the causal susceptibility gene related to this genetic variant.

1.6.1 Functional follow up of candidate SNPs

Even when association results point to a single variant, the true causal SNP is not always easy to identify as the associated SNP can just be serving as a surrogate for the causal one. The association observed with a marker can be a result of an underlying causal allele that is in LD (high r^2) with the associated variant, and is possible that multiple functional alleles are causing the association. The associated SNP can be a common allele meaning that it has high population frequency >1%, or a rare allele meaning with a <1% minor allele frequency.

When a locus with an associated SNP is identified, dense genotyping of markers in that region and performing fine mapping can identify a group of potential causal alleles and the key to this process in order to avoid missing the true causal variant, in case it was not initially genotyped, is to include all known markers in the associated locus. There are two different tests that can be used to analyse the results of fine mapping. One is to apply conditional regression and adjust for the lead candidate causal SNPs in order to remove any other association observed for variants in that region and thus identify the true causal SNP. The second statistical tool, used to identity multiple causal variants is to perform conditional haplotype analysis, where several haplotypes are defined in the hope of finding the haplotype that better explains the risk in that locus. Multiple alleles within *HLA-B* (major histocompatibility complex, class I B) associated with long term viral load control in HIV-infected patients were identified after fine mapping following a GWAS study (Pereyra et al, 2010). It is

important to mention at this stage that when the associated variant is common then fine-mapping is employed to identify the causal SNP, but in cases of a rare SNP, identification is performed using burden testing. This test aims into identifying whether the rare variants interspersed in a locus are distributed in a non-random way suggesting a functional role for them (Bansal et al, 2010).

In the absence of genetic mapping, another method that can be used is to try and translate the SNP's association into possible functional role. If a SNP is a non-synonymous SNP and they induce an amino acid substitution in a protein they might disrupt the function of the protein as previously discussed. A non synonymous SNP that is located in a highly conserved region might be deleterious and several software can be used to predict the function of coding variants (reviewed in Cooper and Shendure, 2011), such as PolyPhen 2 (Polymorphisms phenotypic 2) that predicts the functional importance of an amino acid substitution (Adzhubei et al, 2010).

Assessing the functional role of non-coding variants though is somewhat more challenging. The genome comprises of 99% noncoding sequence and it is not greatly annotated. However, it has been shown that almost 8% of the noncoding genome is under evolutionary constraint, meaning some genome sequences of several species when aligned maintain similarity through evolution, suggesting they are potentially functional sequences (Davydov et al, 2010). These reports are indicating that the noncoding SNPs may have an impact to the functional sequences of the noncoding genome. Some of the approaches used to evaluate the functional role of noncoding sequences are looking at gene expression, sequence conservation mostly used for assessment of rare variants only, chromatin modifications, methylation of CpG islands in gene promoters, or allele specific loss of heterozygosity (Raychaudhuri, 2011). It should be noted here that coding SNPs' functionality may also be investigated by these processes.

Genetic variants can correlate with the transcript level of a gene and are often referred to as expression quantitative trait loci (eQTLs). Most such effects investigated are showing the *cis-effects* of the SNPs that are modulating the activity of a nearby gene's promoter or enhancer or the mRNA stability (SNPs within 1000kb of the relevant gene). *Trans-effects* are SNPs affecting elsewhere

in the genome and have not been largely investigated to date (Freedman et al, 2011). A recent study has identified transcripts whose expression is regulated by SNP-SNP interaction and reported that *cis* acting variants are involved in two-locus gene regulation (Becker et al, 2011). One example of a study that reported both *cis* and *trans* effects of genetic variants on gene expression revealed that differences between ancestral population can be explained by the variation in ancestry affecting the genotype and in turn gene expression (Price et al, 2008). Another genome wide association approach has investigated *cis* effects of SNPs on levels of clinically relevant proteins in human serum and plasma correlating with gene expression (Melzer et al, 2008). A variant upstream of the *KFL14* transcription factor involved in type 2 diabetes has been reported to act in *cis* influencing *KFL14* expression levels and in *trans* influencing many genes regulated by KFL14 (Small et al, 2011). The investigation of the genetic variants' effect on gene expression should be carefully assessed in the right cell models as it is tissue specific. The study that reported tissue specific genotype effect on gene expression has also revealed that common patterns observed across tissues are mainly due to *cis* regulation (Price et al, 2011). Although initially, lymphoblastoid cell lines were used in eQTL studies, more studies have been recently performed in primary human tissues related with the disease studied (reviewed in Freedman et al, 2011).

The investigation of such effects of SNPs associated with EOC risk on gene expression may thus help understand the process of EOC carcinogenesis, by providing hints about the mechanisms for common variants and identifying the likely causal gene. If the gene's transcript is affected by genotype then it might be the causal gene. Increase of that gene's expression with the risk allele may further suggest that the gene increases the risk in EOC development by gain of function, indicating an oncogene. If on the other hand the variant is correlated with decreased gene expression the gene may be involved by loss of function, possibly indicating a tumour suppressor gene (TSG).

1.6.2 Functional follow up of candidate genes

Candidate genes emerging from association studies are genes that are located within associated disease risk loci and could be the target susceptibility genes. These genes could be involved in EOC development, progression and survival. Human *in vitro* models of the cells of origin for the disease could be used as a first line approach to evaluate the function of candidate genes at susceptibility loci identified by GWAS as it is proposed by the aim of this study. When a suitable *in vitro* model is established the role of the candidate gene can be evaluated firstly by evaluating any differential expression between the normal and cancerous primary cells and secondly by further assessing its effect on proliferation, apoptosis, cell cycle progression, anchorage independent growth, migration and chemosensitivity on selected cell lines.

Culturing primary cell lines and establishing the suitable *in vitro* models has proven very challenging. Some examples of these difficulties include the isolation of prostate and colon normal epithelial primary cell lines and the isolation of fallopian tube epithelial cells to study ovarian cancer. One example of a successful *in vitro* model established is 3D cultures of MCF10A normal breast cell lines that closely reflect the architecture of the breast ascini *in vivo*. This model was used for evaluation of EOC susceptibility gene *BRCA1* and loss of its function was correlated with dysfunctional differentiation in the lumen of breast epithelium (Furuta et al, 2005, Proja et al, 2010). In EOC, a study completed in our group has used an IOSE-MYC, IOSE-MYC/KRAS as an EOC transformation model to evaluate the expression of candidate genes that emerged from the ovarian cancer GWAS association study (Song et al, 2009, Goode et al, 2010, Bolton et al, 2010, Lawrenson et al, 2011).

1.7 Somatic genetics of EOC

Ovarian cancer familial cases account for only one in ten ovarian cancer incidences and the rest fall under the category of sporadic cases (no family history). Somatic changes are all the collective differences in the DNA sequence of a cancer cell compared to the germ-line DNA sequence. These alterations can be mutations, insertions or deletions of small or large fragments of DNA, rearrangements, copy number variations where the normal diploid karyotype is increased known as gene amplification and gain or decreased known as copy number reduction resulting in complete absence of a DNA sequence. Somatic changes in turn can lead to altered expression of the genes and deregulation of pathways that lead to carcinogenesis.

Understanding the molecular pathogenesis of EOC's subtypes is important and identification of the genes implicated in its development and progression is crucial for improving its treatment and survival outcome. Initially, it was anticipated that the development of sporadic ovarian cancer would be attributed to a large extent to somatic mutations in *BRCA1*, *BRCA2* and members of the MMR pathway. However, only a small proportion of the sporadic cases have been reported to harbour *BRCA1*, *BRCA2* and *MMR* gene mutations, indicating a difference between hereditary and sporadic ovarian cancer, and ongoing research (Stratton et al, 1997) and constant improvement of the genetic technologies has led to identification of several somatic changes in EOC (The TCGA Study, 2011).

Somatic alterations can lead to the abnormal function of two gene categories and lead to development of cancer including EOC; either by disrupting the activity of tumour suppressor genes or by activating oncogenes. Tumour suppressor genes normally function as inhibitors of cell proliferation. Tumour suppressor genes (TSGs) in cancer act recessively as both copies of the genes have to be inactivated in order for cell proliferation to become excessive. The normal counterparts of oncogenes are the proto-oncogenes and are responsible for the regulation of cell cycle progression, apoptosis, and movement of cells. Oncogenes act in a dominant manner as their activation can be caused by the alteration of a single allele.

A very comprehensive catalogue of the somatic alterations in EOC has been recently published by 'The Cancer Genome Atlas Research Network' (TCGA). TCGA was created as a pilot project in 2006 and was committed to creating a publicly available atlas of the genomic aberrations found in many forms of cancer. The goal of this project was to accelerate the understanding of the molecular basis of cancer and improve diagnosis, treatment and prevention of cancer (The TCGA Study, 2011). An overview of some of the most important mutations identified in ovarian cancer and their clinical implications will be presented.

1.7.1 Common mutations in epithelial ovarian cancer disrupting the normal function of genes

Mutations can be located in the coding regions of genes and cause amino acid changes and can be either activating (in the case of oncogenes) or inactivating (in the case of TSGs).

1.7.1.1 Mutations in tumour suppressor genes and their clinical implications

Inactivating mutations have been found in *TP53, BRCA1, BRCA2, PTEN, CDKN2A* TSGs and some of them have been implicated with the prognosis for EOC patients.

The most common mutation in EOC is found in *TP53* gene which is a TSG frequently found mutated in nearly half of all forms of cancer. p53 has multiple functions including acting as sequence specific transcription factor regulating the activation of other genes involved in a variety of cellular processes and acting as a cellular gatekeeper by monitoring cellular stress and inducing cell cycle arrest and apoptosis to prevent cell growth (reviewed in Bai and Zhu, 2006, Figure 1.7).

Advanced ovarian cancers were reported in early studies to display *TP53* mutations in 40-50% of all cases which are mostly a result of transitions whereas the mutation rate in early stages was only 15% (Berchuck et al, 1994). Later studies that have sequenced the full coding sequence of *TP53* in ovarian

cancers have demonstrated that the prevalence of *TP53* mutation is similar (60-70%) in all stages of serous cases whereas it is significantly lower in endometrioid, clear cell and mucinous cases that are the most common histological subtypes of early stage EOC (Leitao et al, 2004). A recent study has reported the *TP53* mutation signatures to be even slightly higher in early stages, 77% versus 63% in late stages, but yet still comparable. Interestingly, more specific analysis on the type of the mutation (null versus missense) showed significantly higher null (premature chain terminating) mutations in early than late stage of EOC (Bernardini et al, 2010). TSGA has reported *TP53* mutations in 96% of high-grade serous ovarian cancer tumours (The TCGA Study, 2011). Recent research has reported that immunohistochemical analysis of p53 expressing cells in ovarian tumours is a robust method to identify the presence of a *TP53* mutation in ovarian carcinomas (Yemelyanova et al, 2011).

There is controversial data regarding the association between *TP53* mutations and survival. Patients with *TP53* mutations demonstrate 8-fold higher incidence of distant metastasis (Shahin et al, 2000) and significantly worse prognosis (Sood et al, 1999). Others have not found a correlation of null *TP53* mutation occurrence and patient survival but report higher than 2.5 fold presence of missense mutations in short term survivors than long term survivors of EOC (Bernardini et al, 2010). Additionally, EOC patients with *TP53* mutations and/or p53 over-expression have been shown to have worse survival outcome which indicates that loss of functional p53 might confer a chemoresistant phenotype because p53 plays a role in chemotherapy-induced apoptosis (Hall et al, 2004). Again, p53 related prognosis has shown contradicting results favouring the survival and chemosensitivity of patients harbouring *TP53* mutations as reported by other equally valid studies (Lavarino et al, 2000, Havrilesky et al, 2003).

It is possible, that studying and distinguishing between the different types of *TP53* mutations might be a key into understanding the controversial role of p53 in ovarian cancer development. It is conceivable that women with distinct *TP53* mutations which incapacitate p53 in DNA repair are having a more favourable prognosis as a result of their tumour cells inability to repair chemotherapy induced damage. On the other hand, women with other types of

mutations might be less responsive to treatment if their tumour cells are unable to undergo p53-mediated apoptosis.

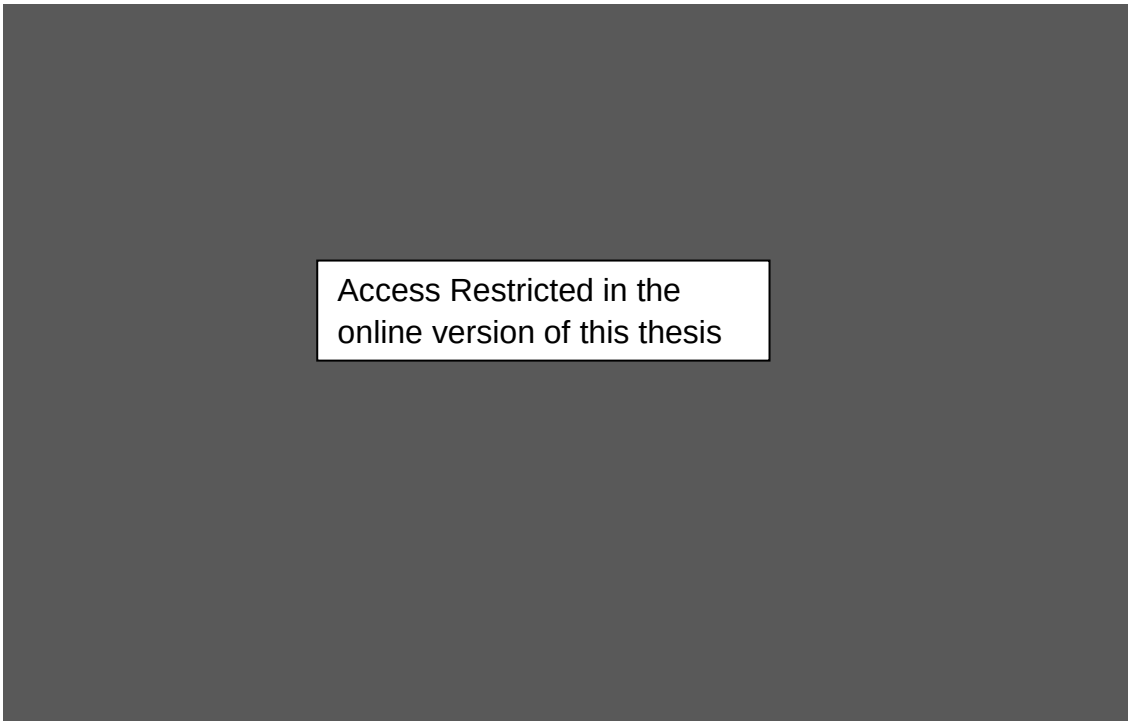


Figure 1.7: Schematic representation of the role of p53 in various stress response pathways. P53 activation from several extracellular and intracellular signal leads to activation of downstream signalling pathways that trigger cell cycle arrest, apoptosis, DNA repair and senescence. (Figure taken from Bai and Zu, 2006).

BRCA1 and *BRCA2* mutations in germline DNA sequence account for <10% of ovarian cancer cases as previously discussed. These are the familial cases and are progressing quicker to late stages but are more responsive to platinum chemotherapy and have a better outcome (Pavelka et al, 2007). *BRCA1* somatic mutations have been documented in 5% or sporadic cases (Stratton et al, 1997) whereas *BRCA2* mutations are relatively uncommon in sporadic ovarian cancer (Foster et al, 1996), but later reports show *BRCA1* and *BRCA2* mutations in nearly 10% of sporadic cases (Jazaeri et al, 2002). TCGA analysis of 316 HGSCs has identified a 3.5% and 3.2% somatic mutation rate for *BRCA1* and *BRCA2* respectively (The TCGA Study, 2011).

The existence of *BRCA1* and *BRCA2* mutations in familial and sporadic ovarian cancer, though in different prevalence, is something that suggests a link between the germline and somatic alteration in EOC carcinogenesis.

Complementary cDNA microarrays were used to investigate the role of *BRCA* mutations in ovarian carcinogenesis by comparing the gene expression patterns in cancers associated with germline *BRCA1* and *BRCA2* mutation and in sporadic ovarian cancers. The findings showed that there is a distinct expression profile difference between tumours with *BRCA1* and *BRCA2* mutations indicating that *BRCA1* and *BRCA2* associated carcinogenesis may arise from distinct molecular pathways. These findings support that there are functional differences between the BRCA proteins, possibly in transcriptional control (Jazaeri et al, 2002).

BRCA proteins are thought to be involved in the DNA double strand break repair pathway (Figure 1.8). Inactivating mutations in these genes confer reduced ability of the cells to repair the DNA damage caused by platinum compounds and lead to death of the cancerous cell. Interestingly, cancerous cells with a defective DNA repair show sensitivity to PARP inhibitors (Drew et al, 2008). In this context somatic *BRCA* mutations might be an important biomarker to personalise treatment and select patient for PARP targeted therapy. Women with high grade serous ovarian cancer harbouring *BRCA* mutations but not BRCA1 deficiency have reported to exhibit improved survival and chemotherapy response (Yang et al, 2011).

Figure 1.8**:** Schematic Illustration of DNA DSB-induced activation of checkpoint and repair pathways. Figure from Bolderson E et al., 2009.

PTEN (Phosphatase and tensin homolog) codes for a protein that acts a cell cycle regulator by inhibiting the survival/growth promoting activity of the phosphoinisitide 3-kinase (PI3K)-Akt signalling pathway and is mapped to chromosome 10q23.3. PTEN blocks Akt activation and indirectly blocks Mdm2 release in the cytoplasm which in turn leads to p53 activation leading to cell cycle arrest and apoptosis.

Somatic *PTEN* mutations are frequently found in endometrioid but not mucinous or serous ovarian cancer and typically more frequently occur in early stage disease indicating that *PTEN* mutations are an early event in ovarian carcinogenesis (Obata et al, 1998), although some studies have identified mutations of *PTEN* in subsets of low grade serous and mucinous tumours (Saito et al, 2000). TCGA analysis of 316 HGSCs has identified a 0.6% somatic mutation rate for *PTEN* (The TCGA Study, 2011).

A reverse association between *PTEN* and *TP53* mutations has been reported in a study that revealed 5% *PTEN* mutations in all subtypes of EOC and 20% in endometrioid alone (Kolasa et al, 2006). Somatic mutations of the *PTEN* gene result in protein inactivation and recurrent somatic mutations are found in CpG dinucleotides (Bonneau et al, 2000).

Finally, *CDKN2A* (Cyclin-dependent Kinase 2A) is involved in cell cycle checkpoint control and was found mutated in approximately 10% of epithelial ovarian cancer (reviewed by Despierre et al, 2010) and TCGA reported that 2.5% of the HGSCs studied harbour somatic mutations (The TCGA Study, 2011).

1.7.1.2 Mutations in Oncogenes and their Clinical implications

Activating mutations have been reported in *EGFR, KRAS, BRAF, PIK3CA, APC, CTNNB1, KIT* and *SMAD4* oncogenes in epithelial ovarian tumours, and the clinical implications of some of these mutations have been studied.

Epidermal growth factor receptor (EGRF) phosphorylation induces activation of the PI3K-Akt and External signal Regulated Kinase (ERK) signalling pathways. *EGFR* over-expression has been linked with poor outcome in ovarian cancer (Skirnisdóttir et al, 2001). Early studies have identified four

EGFR types of mutations reported in ovarian tumours worldwide in exons 19 and 20, but recent mutational analysis in all the exons revealed 27 additional nucleotide substitutions in exons 18, 19 and 20 in Japanese patients which might indicate an ethnic variation in *EGFR* mutations with the mutations present in all subtypes but endometrioid ovarian carcinomas (Tanaka et al, 2011). TCGA analysis of 316 HGSCs revealed a 0.6% somatic mutation rate for *PTEN* (The TCGA Study, 2011).

Data from Non-Small Lung Cancer research have shown that patients with *EGFR* mutations have a better response to the EGFR selective inhibitor Gefitinib but such association was not found in ovarian cancer patients (Schilder et al, 2005). Tanaka et al have demonstrated that the *EGFR* mutational status of those patients did not correlate with EGFR expression, Akt activation or patient survival outcome, however showed that Akt inactivation is correlating with increased sensitivity to platinum based chemotherapy (Tanaka et al, 2011).

The PI3K-Akt pathway is activated by activating hot spot mutations in *PIK3CA* (PI3K catalytic subunit-α) found in a small subset of endometrioid and clear-cell ovarian cancers (Campbell et al, 2004) as well as by somatic *Akt* activating mutations but the latter are a very rare event in ovarian cancer. TCGA analysis of 316 HGSCs has identified a 0.6% somatic mutation rate for *PIK3CA* (The TCGA Study, 2011) but a recent study of 30 ovarian tumours of several subtypes has reported 17% somatic mutation rate for *PIK3CA* (Janku et al, 2011).

Several mutations of members of the RAS-RAF-MEK-ERK-MAP kinase pathway have also been documented in ovarian cancer. The RAS proteins are small GTPases downstream of EGFR and are key components involved in pathways that couple growth factor receptors to mitogenic effectors and influence cell proliferation and differentiation. Several studies have reported *KRAS* activating mutations and collectively the published data support that they are more common in mucinous lesions (45%) than serous (15%) and more prevalent in borderline than invasive tumours with the 10 most common mutations located in codons 12 and 13 of *KRAS* (Fabjani et al, 2005, Auner et al, 2009). These data suggest that *KRAS* is involved in mucinous differentiation. In serous lesions, mutations of *KRAS* and also *BRAF* are typically found in

borderline and low-grade carcinomas. *BRAF* mutations are found in low grade serous ovarian tumours (12%) (Sieben at el, 2004). Consistent with previous reports TCGA analysis has found KRAS or BRAF somatic mutations in only 0.6% of the 316 HGSCs studied (The TCGA Study, 2011).

Mutational status of *KRAS* was not correlated with patients' prognosis (Auner et al, 2009), however bearing in mind that *KRAS* and *BRAF* mutations are indicative of patients response to anti-EGFR treatments in colorectal cancer and are associated with poor survival in colorectal and lung cancer, further work is needed to elucidate their prognostic role in ovarian cancer. In a small study of 30 ovarian tumours 35% were found to have simultaneous *KRAS* or *BRAF* and *PIK3CA* mutations. Patients harbouring simultaneous *PIK3CA* and *KRAS* or *BRAF* mutations were found to respond well to PI3K/AKT/mTOR inhibitors (Janku et al, 2011).

1.7.2 Copy number changes in EOC and their clinical implications

Besides mutational activation of oncogenes and mutations that inactivate TSGs, other frequent events such as copy number variation through chromosomal amplification, gain or deletion can lead to their activation or inactivation respectively. Comparative genomic hybridisation (CGH) and the more high-throughput array CGH (aCGH) have replaced older cytogenetic techniques such as chromosome banding and fluorescence *in situ* hybridisation (FISH) to identify recurrent patterns of copy number gain, amplification, loss (deletion) or allelic imbalance in ovarian tumours.

DNA amplification is a mechanism that allows cells to increase expression of genes that act as oncogenes by regulating cell growth and resistance to chemotherapy. The detection of gene amplification is important for patient disease management for diagnostic, prognostic or therapeutic reasons. Copy number changes have distinct patterns in the different histological subtypes.

MYC amplification in chromosome 8q24 was one of the first copy number changes identified in ovarian cancer (Baker et al, 1990). The same study reported no apparent relationship between *MYC* amplification and response to chemotherapy, which was confirmed in recent studies (Darcy et al, 2009).

Gain of *CCNE1* and *FGF3*/4 was found in all serous carcinomas, and 80% of serous borderline tumours exhibited amplification of *FGFR1* and *MDM2* and 75% showed gain of *PIK3CA*. Previous CGH data have shown gain of *PIK3CA* in 40% of all histological subtypes associated with increased PIK3CA transcription (Shayesteh et al, 1999). Endometrioid carcinomas frequently exhibit gain of *KRAS2* (67%), *MYCN* (50%), *CCND2* (50%), *ESR* and *JUNB* (83%, Mayr et al, 2006). *AKT2* amplification has been reported to 12% of ovarian tumours, and it might be correlated with poor prognosis to paclitaxel as cell lines that exhibited *AKT2* amplification were resistant to paclitaxel (Bellacosa et al, 1995, Page et al, 2000). Amplification of many other chromosomal regions or genes in ovarian cancer has been associated with poor treatment outcome such as *MUC1* in chromosome 1q21, *ERBB2* in 17q21, *CCNE1*, *JUNB* in 19q12 and several others (reviewed in Despierre et al, 2010).

The TCGA Study has recently reported somatic copy number alterations in a study of ~500 HGSCs. The most common focal amplification identified included genes *CCNE1* and *MYC* and each was amplified in more than 20% of the tumours. More tightly localised amplification peaks, showed amplification of a target gene for p53, and the embryonic development gene *PAX8*, and the catalytic subunit of telomerase *TERT*. In an attempt to investigate whether certain gene changes may be linked to response to therapy they identified 22 genes that are therapeutic targets including *KRAS*, *MAPK1* and *CCNE1* that were amplified in 10% of the tested HGSCs (The TCGA study, 2011)

Chromosomal deletions which can be caused by translocations crossovers, inversions or breaks are mechanisms that allow cells to inactivate genes that would normally act as tumour suppressor genes. A few deletions of genetic material in several regions have been reported in ovarian cancer including the 3p, 1p, 17q, 17p13, 4q, 5q, 16q, 17p, 17q, Xp and Xq with the most common abnormality reported in ovarian cancer being the deletion or translocation of chromosome 6q (reviewed in Despierre et al, 2010).

1.7.3 Loss of heterozygosity (LOH) in EOC

Tumour suppressor genes act recessively, which means that the loss or inactivation of both copies of the gene is essential for the normal proliferation of the cells to be rendered uncontrolled. The main strategy to identify tumour suppressor genes has been investigation for LOH or allele loss studies in tumours. Loss of heterozygosity to identify TSGs is based on the two-hit hypothesis formed by Knudson in 1993. The first hit is a mutation in one of the alleles, (inherited in the familial cases, or random in the sporadic) and the second hit may occur by either a mutation or methylation or a deletion of the remaining normal allele. In case of deletion of the normal allele if the patient is heterozygous for a polymorphic marker the tumour will be homozygous, termed loss of heterozygosity (Figure 1.9, B).

Thus, the investigation of multiple polymorphic markers in the tumour DNA compared to germline (blood) DNA of the same individual is used to determine the frequency of LOH in that genomic region. The initial method used to detect LOH was Southern Blot analysis of restriction fragment length polymorphisms (RFLPs) by autoradiography. This was replaced by the amplification of polymorphic microsatellite repeat markers by PCR, initially using radiolabelling and then fluorescent detection. (Cooke et al, 1996, Canzian et al, 1996). SNP arrays now provide high marker density to find regions of LOH. With these SNP arrays, it is possible to identify the absence of the heterozygous loci in a tumour using the hidden Markov model, instead of needing to compare the tumour with a paired normal (Beroukhim et al, 2006).

There are several studies that have reported LOH in sporadic ovarian tumours but two studies have investigated the frequency of LOH on all the chromosome arms in the same panel of tumours, called allelotype studies. The first of the two studies used Southern blotting which is not as accurate and the second one that use microsatellite markers has identified >35% frequency of LOH in chromosomal arms 5q, 6q, 7p, 8p, 9p, 13q, 14q, 17p, 17q, 18q, 21q and 22q (Sato et al, 1991, Cliby et al, 1993). Many chromosomal regions were found to have frequent LOH in malignant epithelial tumours in the last 2 decades, and potential tumour suppressor genes within these chromosomes have been investigated. Candidate TSGs mapped at the arms where the LOH was

detected were subjected to expression analysis in ovarian tumours and/or cells lines to identify a possible role for them in ovarian tumourigenesis.

The events that might lead to LOH are mutation, homozygous gene deletion or epigenetic changes. A frequency of ~20% LOH has been reported in chromosome 1 with the candidate putative TSG *ARHI* identified by differential display PCR, a gene appearing to be involved in growth signal regulation (Peng et al, 2000) through inhibition of STAT3 and FAK-Rho signalling pathways (Badgwell et al, 2011). On chromosome 5q, 46% of allele loss directed to identification of a candidate TSG *Dab2/DOC2* a gene involved in a mitogenic signal transduction pathway (Cliby et al, 1993, Fazili et al, 1999). LOH frequency of up to 65% in chromosome 6q led to identification of the putative TSG *LOT1,* a gene possibly regulating growth signals (Cooke et al, 1996, Abdollani et al, 1999, Niederacher et al, 1999). On chromosome 9, 34% LOH has been found in 9p later possible related with the cell cycle regulatory TSG *CDKN2A* and subsequent studies reported no mutations or deletions or methylation of CDKN2A (Shih et al, 1997) and not significant association of *CDKN2A* gene expression and LOH at its corresponding locus (Niederacher et al, 1999). In chromosome 10q frequency of LOH was reported to be 16% in 10q involving *PTEN.* In chromosome 11p 35% of LOH lead to identification of deletions in the candidate TSG *TSG101* highly expressed in ovarian cancer cell lines (Lu et al, 1997, Liu et al, 2002). High frequency of LOH was reported in chromosome arm 13q involving TSGs *BRCA2* and retinoblastoma 1 (*RB1*), but the low mutation rate reported in *BRCA2* in ovarian cancers indicates other TSG in chromosome 13 might be involved in ovarian tumourigenesis (Sato et al, 1991). The high frequency of LOH (61%) in 17p can partially be represented by alterations found in *TP53* as they are mostly found in advanced stages (Philips et al, 1996). Additionally, methylation analysis revealed *HIC-1* (Hypermethylated in cancer 1 gene) and positional cloning revealed *OVCA1* and *OVCA2* (Ovarian cancer genes 1 and 2) as candidate TSGs implicated in EOC development (Wales et al, 1995, Schrock et al, 1996). Higher than 50% frequency of LOH reported for chromosome 17q is related to *BRCA1* mostly caused by hypermethylation of its promoter region and more rarely by mutation (Wilcox et al, 2005, Despierre et al, 2010). Loss or deletion of chromosome 18q has been

reported to occur in 25 to 40% of ovarian carcinomas (Cliby et al, 1993, Arnold et al, 1996)

1.7.3.1 Allele specific LOH

LOH is non random in TSGs with loss of the normal allele. If LOH at a region was random there would be loss of either of the two alleles. At a population level, random LOH would result in equal loss of either of the two alleles in a large number of tumours. However it is possible that if a SNP has a functional role there may be preferential loss of one allele over the other, as reported in a breast cancer study for LOH in *DAL-1/4.1B*. In this study, a C/T intragenic SNP was analysed for LOH and it was reported that the C allele was preferentially retained, suggesting that allele-specific LOH is occurring (Kanokwan et al, 2004).

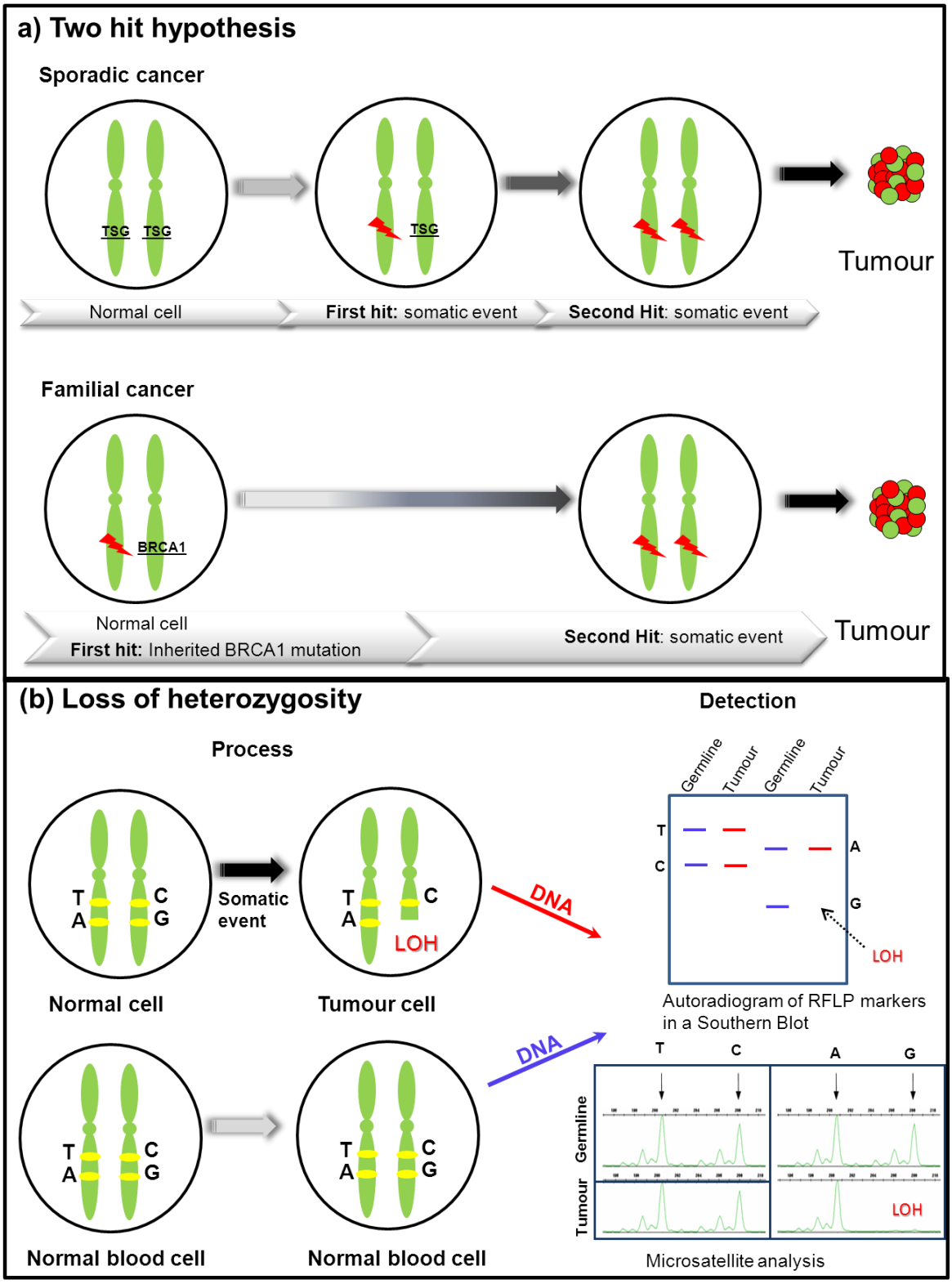


Figure 1.9**:** Two hit Knudson hypothesis and loss of heterozygosity in ovarian cancer. a) Representation of the two hit Knudson's hypothesis for sporadic and hereditary ovarian cancer. The somatic event can be mutation, deletion, or methylation, b) Schematic illustration of loss of heterozygosity and the principle for detecting it using either RFLP analysis or microsatellite analysis.

1.8 Epigenetics

The term "epigenetic" refers to a change in gene expression that is not due to a change in sequence of bases of DNA (Galm et al, 2006). Commonly occurring epigenetic changes are DNA methylation and chromatin remodelling by histone protein acetylation and methylation, as well as nucleosome positioning and non-coding RNAs, specifically miRNA expression. Aberrations in these epigenetic mechanisms have been associated with the development of cancerous lesions and their response to chemotherapy, thus affecting their clinical outcome (Ducasse and Brown, 2006).

1.8.1 Methylation of DNA

DNA methylation is the covalent addition of a methyl group to the 5'-carbon of cytosine in CpG dinucleotide sequences to form the so called fifth base, 5-methylcytosine. The transfer of the methyl group to the cytosine ring from the donor molecule S-adenosyl-L methionine (SAM) is facilitated by DNA methyltransferases (Kumar et al, 1994). The human genome is not methylated on its whole but instead there are scattered methylated and unmethylated segments throughout the genome.

The frequency of CpG dinucleotides in the human genome is 20% less than expected, due to spontaneous deamination of methylated cytosine and conversion to thymine through evolution. From existing CpG dinucleotides scattered around the human genome approximately 60-90% is methylated and the high frequency of their methylation provides a possible justification for the under-representation of CpGs in the human genome. CpG dinucleotide sequences that are found to be very highly concentrated in areas are called CpG islands and comprise 1-2% of the genome (Clark and Melki, 2002). CpG islands appear in lengths of 500-5000 base pairs (bp) and are mostly located at the 5' region of genes and associated with the promoters of more than half of the human genes. In contrast to the bulk of DNA which is methylated and located in regions of gene inactivity, the CpG sites within CpG islands are always unmethylated and are located in regions of genes activity. Absence of

methylation in the CpG islands is a prerequisite for the gene under their control to be actively transcribed and is important for tissue specific genes to be activated (Yoder et al, 1997).

There are two ways via which DNA methylation is believed to regulate gene expression, whether by blocking the binding of transcription factors to the promoter of the gene or by causing the formation of inactive chromatin state (Molloy and Watt, 1990, Bird and Wolffe, 1999).

1.8.2 Methylation in normal development, cancer, EOC and its clinical relevance

The patterns of methylation in the human genome are established in the embryo, initially with hypomethylation of the paternal DNA, following with *de novo* methylation of specific CpG sites just before blastocyst implantation. The methylation state of a CpG island is not always maintained in each cell division and a switch between *de novo* methylation and hypomethylation at each cell division creates a heterogeneous pattern of methylation for any molecule, for example for tissue differentiation where the activation of tissue specific genes is needed, known as epigenetic reprogramming (Reik et al, 2001). However, the general imprinting patterns of methylation occurring in the early embryo are shared and inherited/maintained independently of the heterogeneity that can be found in individual CpG sites. Disruption of these pre-set patterns during adult life has been linked to ageing and disease. The cancer epigenome is characterised by global changes in DNA methylation, which in turn result to altering gene expression profiles and play an important role in oncogenic transformation.

Two types of abnormal methylation in cancer have been identified in cells: hypomethylation and hypermethylation. One possible consequence of hypomethylation of the genome is loss of transcriptional control in normal "silent" areas, which could lead to the expression of genes, such as oncogenes, which are normally disabled (Dunn et al,2003). The second change in DNA methylation observed in cancer is hypermethylation of CpG islands in areas of gene promoters associated with suppression of transcription and is, along with

mutations, a mechanism for inactivating tumour suppressor genes (Herman et al, 2003) and has been proposed as one of the two hits in Knudson's hypothesis for oncogenic transformation. The growing list of hypermethylated and hypomethylated genes in EOC has been extensively reviewed in Barton et al, 2008.

Some of the genes reported to show hypermethylation of the CpG islands within their promoters are genes involved in cell cycle such as *RB1* hypermethylation in glioblastomas (Nakamura et al, 2001). Transcription factors *RAS* homologue *RASSF1A* are also found hypermethylated in most cancers including lung, liver, ovarian and breast cancer (Burbee et al, 2001, Yoon et al, 2001). Genes involved in the repair have been found to be hypermethylated such as *BRCA1* in ovarian cancer (Ibanez et al, 2004). Finally, also genes involved in apoptosis, for example in *TMS1* gene in ovarian cancer have been reported to show hypermethylation (Terasawa et al, 2004).

Epigenetic regulation of *BRCA1* has been extensively studied in ovarian cancer and hypermethylation of its promoter has been found in up to 20% of EOCs, associated with LOH in the *BRCA1* locus and loss of BRCA1 expression. These reports indicate that the second hit for *BRCA1* inactivation may very well be silencing of *BRCA1* expression by hypermethylation (Chan et al, 2002, Ibanez et al, 2004, Teodoridis et al, 2005). Other tumour suppressor genes with reported LOH that have been reported to be down-regulated in ovarian cancer partially as a result of hypermethylation are *ARHI*, *CDKN2A*, *RASSF1A*, *MLH1*, *SFRP1* (a wnt antagonist) and Trail receptor *DR4.* Gene methylation patterns of some genes are also relatively distinct for different histological subtypes or aggressiveness of tumour with examples of a cell cycle regulator inhibitor *SFN* and *WT1* (Wilm's tumour suppressor gene 1) exhibiting hypermethylation mostly in clear cell cancers. *RASSF1* and *APC* (Adenomatous polyposis coli) hypermethylation is more often found in invasive ovarian carcinomas rather than in low malignant potential tumours (reviewed in Barton et al, 2008). Global DNA hypomethylation increases with malignancy in EOCs.

There are much less genes reported to be hypomethylated than hypermethylated. These include metastasis-related gene synuclein-γ (*SNCG*)

and *Sat2* gene that is involved in rearrangement of heterochromatin (reviewed in Barton et al, 2008).

DNA methylation patterns have been also associated with the responsiveness to chemotherapy. Methylated genes implicated in drug resistance are those genes involved in processes known to influence chemosensitivity such as DNA repair and damage response pathways. For example methylation of BRCA1 has been associated with improved response to chemotherapy (Teodoridis et al, 2005). Silencing due to hypermethylation of the DNA mismatch repair gene *MLH1* has been implicated with acquired resistance to platinum based drugs (Plumb et al, 2000).

The identification of epigenetic changes that correlate with tumourigenesis or clinicopathological parameters and patient outcome may reveal new markers that will be used in assessment of risk, therapeutic approaches and individualised treatment.

1.8.3 Genotype-dependent DNA methylation

Evaluating genotype specific methylation could be an additional approach for characterising the potential functional role of SNPs that arise from GWAS. One of the most fascinating discoveries around DNA methylation is the fact that two homologous chromosomes can be differentially methylated. Gametic imprinting and X chromosome inactivation are factors that can cause differential methylation of homologous chromosomes. DNA methylation can be influenced by the environment and chromosome's parental origin and genome sequence. It is possible to investigate DNA methylation patterns in a population along with SNPs typing and see if allelic differences between the chromosomes influence DNA methylation. There are several studies that suggest that the majority of DNA methylation variation is explained by genotype and that genetic variants may have a much bigger impact in DNA methylation than gametic imprinting. A study has reported allele specific methylation of the human *C-HA-RAS-1* gene (Chandler et al, 1987). More recent studies are more powerful in detecting allele specific methylation patterns since recent advances in DNA sequencing technologies.

A family study has recently reported that the different levels between the homologous chromosomes were rarely completely reversed and proposed that DNA methylation is conserved in loci in an allele-specific manner (Gertz et al, 2011). Another recent study linked a SNP in the 5'UTR of *MLH1* with dominantly inherited constitutional epigenetic silencing by studying methylation patterns relatively to genotype in three generations of a family with increased risk in developing cancer (Hitchins et al, 2011). These observations support a model of evolution where genetic variants and not environment, shape epigenetics. Therefore, genetics may be the main regulatory system for oncogenic transformation and epigenetics just a mediator.

1.9 Cell cycle and its regulation

The reproduction of a cell is carried out in a coordinated process that involves four distinct phases comprising the cell cycle, known as G1, G2, S, and M. The basic events that occur during the cell cycle are the accurate duplication of the genetic material taking part during the S phase and the segregation of the replicated chromosomes and cellular organelles into the two identical daughter cells that occurs during mitosis (M phase) (Malumbres and Barbacid 2001). Undifferentiated embryonic cells rapidly perform repeated cycles of DNA replication (S phase) and nuclear and cellular division (M phase). In that case, the replication of the genetic material starts straight after the end of mitosis resulting in the completion of a cell division cycle skipping phases G1 and G2 (Massaque et al, 2004). As embryogenesis further unfolds the needs of the adult cell, being now in a much more complex environment, are increased. As a result, two additional phases take part during the cell cycle, phase G1 and G2. G1 and G2 function as "gap" periods before entry to S or M phase respectively. These allow for the cell to grow and duplicate their organelles as well as providing time for the cell to ensure that the internal and external environmental conditions are optimal for the cell to successfully proceed and complete the S and M phases (Malumbres and Barbacid 2001). On the genomic level, the gap periods are making sure that throughout the cell cycle the genetic material that

is replicated and to be segregated between the daughter cells will be accurate and with no mistakes.

One of the most important features of the cell cycle is its incredible precision, which is achieved by the activity of various checkpoints throughout the cell cycle. At these checkpoints the progress of the cell cycle is monitored, securing the integrity of the genetic material. There are three checkpoints during a cell cycle. One is controlling the transition between G1 and S phase, one is the intra-S phase checkpoint and finally a checkpoint is found between the S and M phase. Although distinct, these checkpoints all respond to DNA damage and share several of the same proteins that sense DNA damage, others that are signal transducers to effector proteins that start a cascade of events that causes DNA repair or in failure of accurate repair, trigger cell cycle arrests and apoptosis (Houtgraaf et al, 2006). Cell cycle checkpoint and chromosome segregation machinery interact with core components of the DNA damage response to ensure that the correct genetic material will be passed on to daughter cells.

1.9.1 DNA damage and mechanisms for DNA repair

The integrity of the genome of eukaryotic cells is constantly challenged throughout the cell cycle. A large amount of damage can be caused to the DNA by exogenous environmental factors such as cigarette smoke, industrial chemicals, UV, several forms of ionizing radiation, and various genotoxic chemical agents. In addition, DNA damage can be induced by endogenous processes such as errors in the DNA replication caused by nucleotide misincorporation and natural by-products created during the normal cell metabolism known as Reactive Oxygen Species (ROS) (Tuteja et al 2001, Houtgraaf et al 2006). There is a variety of DNA lesions that can be caused by these exogenous and endogenous agents including base deletions, mismatches, single-strand breaks (SSBs), double-strand breaks (DSBs), DNA crosslinks and chemical adducts. For example, ionising radiation can cause SSBs and DSBs in the DNA that if not accurately repaired they give rise to mutations and structural rearrangement of the genome (Houtgraaf et al, 2006).

All forms of DNA lesions can lead to DNA alterations and mutations that if not properly repaired are passed onto the progeny and accumulations of such alterations infer a substantial risk of developing uncontrolled cell growth leading to cancer development.

The cell has evolved several complex and sophisticated mechanisms in order to counteract the numerous threats to its genomic integrity and each of those mechanisms is directed to the repair of a specific type of DNA lesion. These mechanisms are part of the cell's DNA Damage Response (DDR) and are distinct pathways, often interlinked, consisting of numerous proteins that form a cascade of events starting from the recognition of the DNA lesion, recruitment of other components of the repair pathway to the site of damage and finally repairing the damage as accurately as possible. The four main DDR mechanisms in a eukaryotic cell are:

Base Excision Repair (BER) pathway: This mechanism is responsible for the removal of destroyed DNA bases. BER is the cell's predominant response to single strand DNA damages that are generated from the cell's metabolism and ROS. Additionally, BER is often involved in the repair of subtle DNA lesions such as SSBs caused by ionising radiation, alkylation products, as well as lesions caused by some chemotherapeutic drugs such as adriamycin and mitomycin (Izumi et al, 2003, Christmann et al, 2003, Houtgraaf et al, 2006). The first step is the detection and removal of the damaged base by recruitment of DNA glycosylases to the site of damage (Scharer et al, 2001). Following the excision of the damaged base the nucleotide gap is correctly filled in by polymerase β (Pol-β) and then ligated by the Ligase3/XRCC1 complex (Houtgraaf et al, 2006). Key enzymes to this process are PARP1 and PARP2 which are acting as sensors and signal transducers for lesions such as SSBs (Lord and Ashworth, 2008).

Nucleotide Excision Repair (NER) pathway: This mechanism is used by the cell for the repair of bulkier adducts on single strand DNA that distort the normal architecture of the DNA helix. Some of these lesions/adducts are pyrimidine dimers that can be caused by ultraviolet light, or adducts that are caused by genotoxic agents such as benzopyrene (Friedberg et al, 2001). NER can be carried out in two distinct pathways, the global genomic repair (GGR)

and the transcription coupled repair (TCR). TCR is taking place when the lesions formed are blocking the path of RNA polymerase during the transcription process and GGR occurs when the architecture of the DNA helix is distorted by incorrect base pairing (Balajee et al, 2000, Mullenders et al, 2001). For both pathways of NER although the detection of the lesion differs, the repair process is identical with the excision of the DNA around the lesion and the gap repaired by DNA polymerases and finally ligated with ligase I. Important enzymes involved in the NER pathway are XPA that confirms the existence of the DNA lesions after the DNA helix is unwound and ready to be excised and the Excision repair cross-complementing proteins 1 and 4 (ERCC1 & ERCC4) which are the key proteins during the DNA excision step (Houtgraaf et al, 2006).

Mismatch Repair (MMR) pathway: This repair mechanism is responsible for the removal of non-complementary bases that occur either in the process of deamination, oxidation or methylation of bases or by transcriptional errors (Umar et al, 1996). The recognition of the modified bases is carried out by the MutSα complex, consisting of homologous proteins MSH2 and MSH6. MSH2 can also be couple with the repair protein GMSH3 forming the complex MutSβ. Each complex has different substrates and thus plays a different role in the MRR process (Christmann et al, 2003). Once the MutSα complex recognises and binds to the mismatched base, it recruits MLH1 and PMS2 homologous repair proteins to the complex. The excision of the DNA sequence that contains the mismatched base is carried out by exonuclease I and the correct sequence is filled in by polymerase δ (Longley et al, 1997, Genschel et al, 2002).

Double strand break (DSB) repair mechanism: DNA DSBs are the most threatening DNA damages triggered in response to genotoxic stress and can lead to chromosomal abnormalities such as fragmentations, deletions and translocations and further to cell death. Thus, the rapid repair of such DNA damages is one of extreme biological significance. There are two mechanisms for dealing with the repair of DNA DSBs: **homologous recombination (HR)** and **non-homologous end joining (NHEJ).** The choice of which mechanism to be used is largely dependent upon the cell cycle phase when the DSB occurs. HR mainly takes place during S or G2 phases. For HR to occur, necessary requirements that have to be fulfilled are the DNA replication and the two sister

chromatids formation. In contrast, NHEJ can take place throughout the cell cycle but mainly occurs during the G0/G1 phases (Takata et al, 1998). HR is a complex conservative mechanism restoring the exact DNA sequence that was damaged. NHEJ is simpler but is error prone with the result of loss of genetic information and mutagenesis.

During **non homologous end joining,** the ends of the broken DNA strands are joined, without necessarily requiring DNA homology at the points of the joining. The first step in this process is the formation of a heterodimer complex, consisting of the proteins Ku78 and Ku80, which protects the damaged area from being cleaved by exonucleases. After binding to the DNA, the heterodimer interacts with the catalytic subunit of DNA-PK (XRCC7, DNAPKcs) forming a holoenzyme. The DNAPKcs complex activates XRCC4 which binds to ligase IV, which is turn joins the broken ends. Before the joining of the broken ends, damaged DNA on the 3' end is removed by the complex MRE11-Rad50-NBSg1 which have an exonuclease, endonuclease and helicase role respectively and DNA is removed from the 5' end by endonuclease FEN 1(Critchlow et al, 1998, Christmann et al, 2003).

During **homologous recombination,** DNA homology is a requirement and the damaged chromosome is in direct contact with a homologous intact DNA sequence from the sister chromatid that is acting as a template for the synthesis of new DNA at the DSB site. The most important proteins involved in HR are BRCA1, BRCA2, PALB2, ATM, CHK1, CHK2, Rad51 and Rad52 (Houtgraaf et al, 2006). In the event of a DNA DSB, a complex formed of BRCA1-abraxas-Rap80 binds to the ubiquitylated histones following phosphorylation of the histone gamma H2AX (γH2AX). The damage sensor complex formed by MRE11-Rad50-NBS1 (MRN complex) is processing the ends of the two broken DNA chains in a 5'-3' direction. The BRCA1-CtIP complex is in turn associated with the MRN complex and through ATM and CHK2 dependent phosphorylation of BRCA1 the BRCA1-PALB2-BRCA2 complex is formed to trigger the Rad51 mediated HR (Roy et al, 2012). The protein Rad51 searches the genome for an intact copy of the damaged DNA on the sister chromatid. Each damaged 3' DNA strand is joined with a heptamer complex that is formed by Rad52 interacting proteins, which protects it from further cleavage by exonucleases.

The Rad52 complex interacts with the Rad51 complex (Rad51C, Rad51D, XRCC2 and XRCC3) activating the process of exchange of homologous parts of genetic material between the damaged and the intact sister chromatid. Another important protein that interacts with Rad51 is RPA, which acts to stabilise the Rad51 inducible exchange of homologous DNA parts. After completion of the DNA synthesis, the final ligation and branch migration of the repaired strands follows. (Golub et al, 1998, Stasiak et al, 2000, Sonoda et al, 2001, Christmann et al, 2003, Houtgraaf et al, 2006, Roy et al, 2012).

1.9.2 Defects in DNA damage response in neoplastic development

As previously discussed, tumours are often characterised with genomic and chromosomal instability, and DNA alterations such as nucleotide substitutions, insertions and deletions. Defects of the DDR have often been linked to neoplastic development by giving rise to those DNA abnormalities. One example is mutations in the mismatch repair genes *MSH2* and *MLH1* are linked to hereditary non-polyposis colorectal cancer. Another example are silencing mutations or epigenetic silencing of *BRCA1* and *BRCA2* genes causing defects in homologous recombination that is found defective to 50% of high-grade serous ovarian adenocarcinomas both in familial and sporadic ovarian cancer.

There is substantial evidence that DDR is dysfunctional during neoplastic development. DNA DSBs marker, γH2AX nuclear foci formation, has been found to be markedly elevated in precancerous lesions. There is the hypothesis that activation of oncogenes such as *MYC* and *RAS* causes the formation of multiple replication forks that stall and collapse triggering the DDR and cell cycle checkpoint to repair DNA lesions before mitosis takes phase. Inactivation of ATM, ATR and p53 which play a critical role in DNA DSB repair and cell cycle checkpoint control can in turn cause the progression to mature tumours from those precancerous lesions. With important players of the DDR becoming inactivated the cells proceed through the cell cycle with unrepaired DNA lesions increasing the chance of oncogenesis (Martin et al, 2008).

A cancerous cell is using the normal mechanisms of the cell for further growth and survival. Thus, several cancer treatments are to date aiming in exploiting DDR defects in order to target and eliminate the cancer cells. For example, platinum agents that cause inter and intrastrand crosslinks in DNA are effectively used in the treatment of high grade serous ovarian cancer since it has been reported that HR is often defective in such cases due to inactivation of the *BRCA1* and *BRCA2* genes. Similarly, platinum therapy is effective in patients that have *ERCC1* mutations and thus defective NER pathway responsible for removal of large DNA adducts. Recent evidence have shown that a promising approach is to target directly specific components of a DDR pathway such as topoisomerase inhibitors that cause DNA breaks across the genome to be unrepaired (Martin et al, 2008).

Very successful specific inhibitors are PARP inhibitors (Lord and Ashworth, 2008). Members of the PARP family are involved in DNA repair through three different mechanisms and thus are a very attractive target for therapy. Firstly, though direct interaction of PARP1 with XRCC1, Polβ and PARP2 interacts with XRCC1, Polβ, DNA ligase III that are key proteins in the BER pathway (Masson et al, 1998, Schreiber et al, 2002). Secondly, PARP1 is modifying chromatin structure, after DNA damage occurs. PARP1 binds to the 20S proteasome and increases its proteolytic activity. This leads to the degradation of oxidatively damaged histones and chromatin structure modulation with in turn facilitates the access of DNA repair enzymes to the site of DNA damage (Mayer-Kuckuk et al, 1999, Ullrich et al,1999). Finally, a specific binding site on members of the PARP family enables their interaction with a large number of repair enzymes and molecules involved in the checkpoints of the cell cycle, such as P53, P21, XPA, MSH6, DNA ligase II, XRCC1, DNAPKcs, Ku70, NF-kB. With their poly ADP ribosylation PARP proteins can regulate DNA repair, cell cycle regulation and apoptosis (Pleschke et al, 2000).

1.10 Survival and Chemoresistance in EOC_An overview

The reported overall survival rates for stage II-IV diagnosed patients is 19-30% (Pfeiderer, 1984, Gonzalez-Diego et al, 2000). For EOC treatment platinum therapy is the most widely used which includes cisplatin or carboplatin, often in combination with taxane therapy (paclitaxel) (Rustin et al, 1996). Although chemotherapy is widely used to treat EOC the survival rate hasn't improved over the last decades due to disease recurrence, possibly caused by drug resistance that has a significant role in tumour progression (Bradshaw and Arceci, 1998, Haq and Zanke, 1998).

Some of the known mechanisms for platinum resistance include reduced drug delivery to targeted DNA and also increased ability of cells to respond efficiently to DNA damage induced by platinum and defected genes in the DNA repair, apoptosis and cell cycle control pathways (Kelland et al, 2007). There are many studies that show that germline genetic variation has a role in survival and further in chemoresistance. One study has shown that polymorphisms in glutathione S-transferases (*GSTs*) can be a valuable marker to determine patient prognosis as they found significantly better survival in advanced EOC patients with the null *GSTM1* genotype (Medeiros et al, 2003). Polymorphisms in the epidermal growth factor (*EGF*) also have been reported to be associated with EOC survival (Heffler et al, 2007). Another study has shown that patients with amplifications in Cyclin E (*CCNE1*) gene had poor survival and short response time (Etemadmoghadam et al, 2009). Several studies have also shown that there is a significant survival difference between EOC patients with and without *BRCA1* and *BRCA2* mutations (Chetrit et al, 2008, Tan et al, 2008). Survival analyses have been performed in studies to identify molecular markers that affect survival and this could possibly be related to chemoresistance. This could lead to better understanding of the causes of chemoresistance in groups of patients and more efficient individualised treatment but also the development of valuable tools for accurate prognosis. Thus, a lot of research focuses on using *in vitro* cell models to provide information that will enable the development of treatment that can overcome drug resistance.

Several studies have shown that p53 function is altered in drug resistant EOC cells and that when lost it confers resistance to cisplatin (Fajac et al, 1996)

whereas sensitivity to paclitaxel is retained regardless of *TP53* status (Jones et al, 1998). Another study has reported a mechanism of resistance to paclitaxel by showing that the up-regulation of the transcription factor *NAC-1* confers resistance to paclitaxel (Jinawath et al, 2009). DNA methylation has also been associated with chemoresistance when after treating EOC cell lines with the hypomethylation agent azacitidine, increased sensitivity to carboplatin was observed and this mechanism was further linked to the induction of apoptosis through the caspase 3 and caspase 8 pathways (Li et al, 2009). An increased efflux of cisplatin has been reported in some cisplatin resistant cell lines and it has been proposed that cisplatin might be exported from the cells using chaperones and transporters that mediate copper homeostasis (Katano et al, 2002). Other studies have tried to silence genes in order to evaluate their involvement in drug resistance, for example a study reported that knocked down clusterin in cisplatin resistant cell lines can lead to increased resistance to cisplatin and inhibited proliferation, invasion and migration (Wei et al, 2009).

These studies and numerous others have shown that differential gene expression, protein conformation and function, methylation patterns, impaired uptake of platinum compounds are some of the mechanisms that can lead to resistance to chemotherapy. There is now a growing list of proposed mechanisms and molecular markers of chemoresistance in hope of improvement of individualised therapy.

1.11 Summary

Epithelial ovarian cancer (EOC) is a complex disease that is characterised with molecular and histological heterogeneity and thus should not be treated as a single disease. The heterogeneity of EOC serves as a major challenge for elucidating its initiation, development and progression and the appropriate *in vitro* models should be selected depending on the subtype to be studied. Both NOSE and FT *in vitro* models could be used; and recapitulating in the most appropriate way the tissue micro-environment could be achieved by establishing three-dimensional models of the appropriate cell of origin for the subtype to be studied.

The different subtypes of EOC can be classified not only according to the proposed site of origin, but also distinct DNA alterations. EOC is a polygenic disease originating from different sites and can be familial or sporadic. Although the high susceptibility genes *BRCA1*, *BRCA2*, and gene members of the MMR pathway are responsible for approximately half of the familial incidence of EOC, a large proportion of familial cases as well as the reported sporadic cases accounting for 90% of EOC, are not attributed to defects in those genes. The remaining susceptibility genes could be of moderate or low penetrance and can be identified by genetic association studies which aim to discover associations between polymorphisms and cancer susceptibility by genotyping SNPs in a population. Genetic association studies initially investigated potentially functional SNPs, such as non-synonymous coding SNPs, or focused in polymorphisms tagged in candidate genes of pathways implicated with the disease. Genome wide association studies (GWAS) have emerged in the last decade as a powerful tool for evaluating the frequency of thousands of SNPs in a genome wide manner regardless of an existing proposed function for these variants. GWAS are performed using a staged study design where significant associations are taken to the next stage to be genotyped in larger sample numbers. The most significant result of a GWAS to date has been the identification of an association of a SNP within FGFR2 with increased risk for breast cancer ($P= 2\times10^{-76}$). An ovarian cancer GWAS performed by our group and collaborators has reported significant associations of SNPs in loci 2q31, 3q25, 8q2, 9p22, 17q21 and 19p13.

The main challenge emerging form the vast amount of data reported by GWAS has been to translate the implication of the risk (or survival) associated loci in disease development by investigating how the candidate risk associated alleles affect risk modification and which are the causal susceptibility genes related to these risk variants. Firstly, it should be noted that the initially genotyped risk associated SNP might be the putative causal variant but could be correlated with the actual causal variant(s) which can be determined by performing fine mapping of the associated loci.

Several attempts have been made to assign functional relevance of risk associated SNPs emerging from GWAS. Bioinformatics *in silico* analyses can be performed to identify any predicted function for such SNPs which may be causing amino-acid substitutions in case of coding SNPs or in the case of non-coding SNPs their potential impact on the functional sequences of regulatory elements for gene expression, promoter methylation and chromatin modifications. Some of the approaches that can be used to evaluate *cis* and/or *trans* effects of SNPs in transcriptional regulation include investigating gene expression, methylation of CpG islands in gene promoters, chromatin modifications, miRNA expression, gene splicing and loss of heterozygosity in respects with genotype using *in vitro* assays. Examples include the reported association of a risk allele with increased expression of the *PVT1* oncogene in locus 8q24 (Meyer et al, 2011), and increase of *FGFR2* expression in the rare homozygote samples of the risk associated SNP for breast cancer using microarray analysis (Meyer et al, 2008). However, no studies have proposed yet a systematic approach investigating several aspects of possible functional effects or GWAS derived risk associated alleles as well as the functional role of candidate genes within the identified loci.

A systematic functional investigation of GWAS identified risk or survival associated loci using appropriate tissue specific *in vitro* models will elucidate the way in which low-moderate risk variants lead to the initiation of EOC pathogenesis and will be beneficial for the more effective screening prevention and individualised treatment of the disease.

1.12 Aims of the thesis

The overall aim of this thesis was to employ appropriate models to investigate the functional role of genetic variants and candidate genes of low/ moderate effect that emerge from pathway based approaches and GWAS in epithelial ovarian cancer.

Given the recent proposal for multiple cells of origin for EOC, it was hypothesised that it is necessary to study EOC tumorigenesis using both NOSE and FTE *in vitro* models. Chapter 3 describes the establishment and characterisation of FTE cell lines and tests the hypothesis of whether 3D models of FTE cultures more closely resemble the *in vivo* characteristics of the FTE than traditional 2D cultures. Additionally, the largest primary NOSE cell line repository was established and characterised as it was hypothesised that studying differential expression in NOSE & FTE compared to EOC cell lines would be a powerful tool to identify whether candidate genes are implicated in EOC development.

The aim for the work described in chapter 4 was to test whether variants within candidate genes selected from a pathway based approach may be implicated in EOC development. The potential functional role of candidate SNPs was evaluated by assaying allele specific LOH in primary tumours. Additionally, the clinical implication of the candidate genes was evaluated by testing whether LOH in these genes would affect the survival of EOC patients.

In Chapter 5 the post-GWAS functional characterisation of EOC risk associated loci is described. Although the ovarian cancer GWAS allowed the discovery of genetic variants significantly associated with ovarian cancer susceptibility, it is not known whether candidate genes within the associated loci region are the susceptibility targets or if they have a somatic role in EOC. Additionally, is it is unknown whether the associated SNPs have a regulatory role in EOC development. The functional role of candidate genes within a 1Mb region of EOC susceptibility SNP was investigated in the established NOSE & FTE versus EOC cell lines expression model. Furthermore, in order to investigate the potential synergy between genetic variation and epigenetic or gene expression changes into the development of sporadic EOC, genotype

specific methylation and genotype specific gene expression in non-tumour samples were performed.

Even if post-GWAS characterisation studies identify candidate genes that may have a role in EOC, further functional assays in appropriate *in vitro* models should be performed in order to elucidate the molecular mechanisms by which these genes are involved in EOC development. Therefore, in chapter 6, *MERIT40*, one of the GWAS candidate genes found to have differential expression between normal (NOSE & FTE) and EOC cell lines, and an intriguing function as a BRCA1 associated protein in DNA repair, was knocked down in EOC cell lines and reversal of neoplastic phenotype and re-sensitisation to platinum drugs was assayed.

2 Materials and Methods

2.1 General cell culture

2.1.1 Plasticware and general cell culture equipment used

Cell culturing was performed in a laminar-flow safety cabinet (Heraus), and cells were incubated in a Heracell incubator (Heraus) at 37^{o}C, 5% CO_2. The cells were grown in tissue culture treated disposable 60mm, 100mm and 145mm dishes, 25cm^2, 75cm^2 and 175cm^2 flasks, or 6, 24 and 96multiwell plates (Greiner, Corning or Appleton Woods Ltd). Falcon tubes of 15ml or 50ml (Greiner) were used to centrifuge cell suspensions. Cell culture pipette tips 1ml, 5ml, 10ml, 25ml were purchased form Greiner or Appleton woods. All glassware was washed and autoclaved by heating to 121^{o}C for 15 minutes at a pressure of 103kPa to sterilise.

2.1.2 General tissue culture techniques and common reagents

All tissue culture media were stored at 4^{o}C. Antibiotic and growth factor solutions were stored at −20^{o}C. Puromycin solutions (Sigma) were prepared in ethanol (Sigma). Gentamycin solutions were purchased form Sigma. Penicillin/Streptomycin 100× was provided by Gibco. Cells were washed with sterile 1X phosphate buffered saline (PBS) (VWR). A 10× PBS solution was diluted 1:10 with DDW and the 1× solution was sterilised by autoclaving.

Foetal bovine serum (FBS) (Invitrogen or Lonza) was batch tested before use. For this normal ovarian fibroblasts were used that are sensitive to serum. Cells were mass cultured in the test serum for a minimum of 7 days before plating at clonal density (Three replicates of 100 cells per well in a 6 well plate). Cultures were re-fed every 2-3 days and stained with crystal violet. To stain cells, cultures were fixed in 100% methanol (VWR) for 10 minutes and then rinsed with DDW. Cultures were then stained with 0.5% crystal violet (Sigma) solution (in DDW). After 10 minutes plates were rinsed with DDW and left to air

dry. Colonies were then counted. Colony formation efficiency (CFE) was calculated using the following formula:

$$\frac{\text{Total Number of Colonies}}{\text{Initial Number of Cell Plated}}$$

2.1.3 Passaging, counting and freezing cells

To passage the cells, the cell monolayer was washed with sterile PBS, and 500µl-1ml of 1x Trypsin-EDTA (Invitrogen) was added to the flask and incubated at 37°C for 2-5 minutes for the cells to detach. Trypsin was then inactivated with 10x higher volume of media containing FBS. The cells were resuspended gently to break clumps and reseeded to the required dish or flask.
To freeze cells, following cell detachment and inactivation of Trypsin, the cell suspension was collected and centrifuged at 1,500rpm for 5 minutes to pellet cells. The supernatant was then aspirated and the cell pellet resuspended in freezing solution for cryopreservation. Freezing solution was prepared using 10% DMSO (Sigma) in FBS (Invitrogen). Resuspended cells in freezing solution were aliquotted into freezing vials (Nunc) and the vials were stored in a Mr Frosty container overnight at -80°C, to slowly cool at 1°C per minute before cells were transferred into liquid nitrogen for storage at −196°C.

For cell counting, following cell detachment and inactivation of Trypsin, cells were resuspended gently to break cell clumps and counted under the light microscope using a haemocytometer (Neubauer).

2.1.4 Mycoplasma testing

All cell cultures were tested frequently for mycoplasma contamination. When cell cultures are infected with mycoplasma it is not visible like other contaminations as no turbidity is caused in the media and they are resistant to antibiotics routinely used in cell cultures. Mycoplasma infections if unnoticed can affect gene expression and morphology of cells. Two methods were used to detect mycoplasma infections, Hoescht staining and polymerase chain reaction (PCR). Hoescht fluorescently stains mycoplasma DNA which makes it then visible under a fluorescent microscope. Hoescht staining has a sensitivity of detecting mycoplasma of a concentration of least 106cfu/ml. For this

technique cells had to be incubated for two weeks in antibiotic free media to prevent possible antibiotic suppression of mycoplasma so lower levels than could be detected. The cells were then split on dishes containing coverslips and when they reached 60% confluency they were fixed in freshly-made Carnoy's fixative (3:1 methanol:acetic acid, Sigma). For this, 3ml of fixing solution was added to the cell culture media and incubated for 5 minutes in room temperature. Fixative was discarded and 5ml of fresh fixative was added for 5 more minutes. The last step was repeated and the coverslips were then washed with DDW and air dried. 50µl Hoescht bisbenzimide solution (0.05µg/ml in DDW, Sigma) was then added to coverslips, and cells were observed under the fluorescent microscope. Mycoplasma positive cells were identified by bright spotty cytoplasmic staining (mycoplasma DNA). In cultures that are free of contaminating mycoplasma, only the cell nuclei should be detected. For analysis of staining, positive and negative controls were stained in parallel.

For mycoplasma screening by PCR, 1ml medium was removed from the cell cultures. The aliquoted media was spun at 13,000rpm for 30 minutes, and the majority of the supernatant was then removed leaving approximately 20µl. Pellets (not visible) were then resuspended in 50µl of TE buffer (10mM Tris-HCl with 1mM EDTA) with added Proteinase K (200µg/ml final concentration, Sigma), and incubated for 1h at 55°C. Proteinase K was deactivated by incubating at 98 − 100°C for 10 minutes. PCR was performed using the following primers that detect all mycoplasmas targeting the 16S rRNA gene:

Forward primer (MYCO-F): 5'-GGGAGCAAACAGGATTAGATACCCT-3'

Reverse primer (MYCO-R): 5'-TGCACCATCTGTCACTCTGTTAACCTC-3'

The PCR reactions were prepared as follows:

Reagent	**Volume added (µl)**
dH_2O	7.38
10X PCR Buffer (Applied Biosystems)	1.5
25mM MgCl (Applied Biosystems)	1.5
10mM dNTPs (Invitrogen)	1.5
MYCO _F primer	1
MYCO _R primer	1
AmpliTaq Gold (Applied Biosystems)	0.12
Sample	1
Total volume	14

The PCR was performed using the 9700 PCR machine at 55^{o}C annealing temperature, 40 cycles. Then the PCR product was subjected to electrophoresis on a 1.5% agarose gel. The 1.5% agarose (Sigma) gel was prepared in 1xTBE buffer with the addition of ethidium bromide. 5µl of 1Kb DNA Ladder (Promega) was added on the first lane of the gel. 2µl of 5x DNA loading dye was added in 8µl of PCR product and then loaded on the gel. The gel was then run at 90V for ~1hour 15 minutes. The gel was visualised under Ultraviolet (UV) light where PCR product was confirmed.

2.1.5 Cell lines used in this study and their culturing conditions

2.1.5.1 EOC cell lines

All EOC cell lines of several subtypes were cultured in the following media listed in Table 2.1 and incubated at 37^{o}C, 5% CO_2. All the media and reagents were purchased form Sigma apart from L-Glutamine which was purchased from Invitrogen and FBS which was purchased from GIBCO.

EOC cell line	Subtype	Culturing media
A2780	Serous	RPMI 1640 +10% FBS + L-Glutamine
A2780CP	Serous	DMEM:F12, 10% FBS
COV318	Serous	DMEM +10% FBS+L-Glutamine+Lasparagine
FUOV1	Serous	DMEM/F12 + 10% FBS
HEY	Serous	RPMI 1640 +10% FBS + L-Glutamine
HEY A8	Serous	RPMI 1640 +10% FBS
MPSC1	Serous	RPMI 1640 +10% FBS
OAW 42	Serous	DMEM + 10% FBS + Na Pyr + 20µg/ml Insulin + L-Glutamine
OAW 42M	Serous	DMEM + 10% FBS + Na Pyr + 20µg/ml Insulin + L-Glutamine
OVCA 429	Serous	MEM EAGLE + 10% FBS + NEAA + L-Glutamine + Na Pyr
OVCA 433	Serous	MEM EAGLE + 10% FBS
OVCAR 5	Serous	RPMI 1640 +10% FBS + L-Glutamine
OVCAR-3	Serous	RPMI 1640 +10% FBS
OVM215	Serous	RPMI 1640 +10% FBS
PEO-14	Serous	RPMI 1640 +10% FBS + L-Glutamine
SKOV 3	Serous	RPMI 1640 +10% FBS
SKOV 3IP	Serous	RPMI 1640 +10% FBS
UWB1.289	Serous	RPMI 1640 +10% FBS
UWB1.289+BRCA1	Serous	RPMI 1640 +10% FBS
1847	Unknown histology*	RPMI 1640 +10% FBS + L-Glutamine
1847 AD	Unknown histology*	RPMI 1640 +10% FBS + L-Glutamine
CAOV3	Unknown histology*	DMEM + 10% FBS + NEAA/glucose
COV 413	Unknown histology*	DMEM +10% FBS+L-Glutamine+Lasparagine

COV 624	Unknown histology*	DMEM +10% FBS+L-Glutamine+Lasparagine
DOV 13	Unknown histology*	MEM, 10% FBS + NEAA (Non essential aminoacids)
HOC 7	Unknown histology*	DMEM + 10%FBS
IGROV1	Unknown histology*	DMEM + 10%FBS
INTOV-2	Unknown histology*	RPMI 1640+ 10% FBS+ L-Glutamine+ 50uL β Mercaptoethanol +1% Na Pyr
JAMA-2	Unknown histology*	RPMI 1640 +10% FBS + L-Glutamine
LK1	Unknown histology*	RPMI 1640 +10% FBS + L-Glutamine
LK2	Unknown histology*	RPMI 1640 +10% FBS + L-Glutamine
OAW-41M	Unknown histology*	RPMI 1640 +10% FBS + L-Glutamine
OC316	Unknown histology*	RPMI+10%FBS+L-Glutamine
OVCA X	Unknown histology*	RPMI 1640 +10% FBS
OVCAR 8	Unknown histology*	RPMI 1640 +10% FBS + L-Glutamine
OVCAR-10	Unknown histology*	RPMI 1640 +10% FBS
OVP1	Unknown histology*	RPMI 1640 +10% FBS + L-Glutamine
PXN94	Unknown histology*	RPMI 1640 +10% FBS + L-Glutamine
COV 644	Mucinous	DMEM +10% FBS+L-Glutamine+Lasparagine
EFO 27	Mucinous	RPMI 1640 + 20% FBS + L-Glutamine + NEAA + Na Pyr
COV-434	Granulosa cell	DMEM +10% FBS+L-Glutamine+Lasparagine
C13	Endometrioid	DMEM +10% FBS+L-Glutamine+Lasparagine
OV2008	Endometrioid	RPMI 1640 +10% FBS + L-Glutamine
OV-TRL-90T	Endometrioid	M199:MCDB105 + 15% FBS+ L-Glutamine
TOV 112D	Endometrioid	M199:MCDB105 + 15% FBS+ L-Glutamine
ES-2	Clear cell	McCoy's 5A + 10%FBS + L-Glutamine
TOV 21G	Clear cell	M199:MCDB105 + 15% FBS+ L-Glutamine

Table 2.1**:** Culturing media of EOC cell lines used in this study. * the collection of these cell lines spans over several years and the terminology of histopathological examination has varied. The term “unknown histology” is referring to cell lines that their subtypes is probably serous but collected long ago before the subtypes were well differentiated from pathologists.

2.1.5.2 Normal ovarian fibroblast and non-ovarian cell lines

HUVEC, T71 and INOF2 were used later as controls for immunofluorescence staining experiments and 293T was used to package lentiviruses. The culturing conditions of these cell lines are shown in Table 2.2.

Cell line	Type	Culturing media
HUVEC	human endothelial from umbilical cord	EBM media (Cambrex Cat No CC-3156)
293T Cells	human embryonic Kidney/ retroviral producing	DMEM, 15% FCS, L-Glutamine
INOF2	a normal ovarian fibroblastic cell line	DMEM, 15% FCS, L-Glutamine
T71	endometrial cancer cell line	MCDN105:Media199, 15% FCS, L-Glutamine

Table 2.2**:** Culturing media of normal ovarian fibroblast and non-ovarian cell lines used in this study.

2.2 Establishing and characterising primary cell lines

2.2.1 Collection and culturing of Normal Ovarian Surface Epithelial (NOSE) and Fallopian Tube Epithelial (FTE) cell lines.

NOSE cell lines: The dates of operations for eligible patients for collection were confirmed on the operating schedule list in University College London Hospital (UCLH). Patients that were suitable for NOSE cells collection were patients that were subjected to Total Abdominal/Laparoscopic Hysterectomy (TAH/LH) due to endometrial cancer or endometriosis and there were no suspicion of metastasis to the tubes or ovaries. Patients' consents were collected by a trained nurse the day before surgery (Mrs Nyaladzi Balogun). All informed consents from donors were prepared according to a protocol approved by the Joint UCL/UCLH Committees on the Ethics of Human Research (Committee Alpha). During surgery, when the specimen was removed from the patient it was placed on a sterile kidney dish. Using sterile gloves the uterus was orientated so that the flat side was at the bottom. Each ovary was then gently brushed twice with a sterile Transwab brush swab (Medical Wire and Equipment Ltd) and the cells were placed in a tube containing NOSE culturing media (NOSE-CM). NOSE-CM was composed of MCDB 105 media pH 6.5 (containing L-Glutamine & 25mM Hepes), (SIGMA) and Media 199, (SIGMA) in a 1:1 ratio. and supplemented with 15% FBS, 10ng/ml Human recombinant Epidermal Growth Factor (hEGF) (Invitrogen), 0.5µg/ml Hydrocortisone (HC) (SIGMA), 5µl/ml Human recombinant insulin (GIBCO or SIGMA), 35µg/ml Bovine Pituitary Extract (BPE) (GIBCO) and 50µg/ml Gentamycin sulphate salt (SIGMA). The cells were placed in the tube containing NOSE-CM by shaking the cytobrush. NOSE cells were then cultured in a T25 flask at 37°C, 5% CO_2 until 50-70% confluent. The medium was replaced with fresh every 3 days. Cells were then trypsinised as described above, and split in a ratio of 1:3-1:7, depending on the growth characteristics of each individual cell line. No antibiotics were used after the second passage. A NOSE primary cell line collection database was created where the date of collection, left or right ovary collected, diagnosis and surgical procedure performed, patient number and date of birth were recorded. Samples were kept only when confirmed free

of tumour by the pathology report. A pathology report was generated stating the exact stage and form of endometrial cancer the patient had and confirming if the ovaries were normal and free of tumour (Dr Rupali, Dr Benjamin).

FTE cell lines: A similar procedure than the one used for NOSE was adopted with the difference in the sample collection procedure during surgery. In summary, the ampullary region of the FT of women that underwent total abdominal hysterectomies but had normal tubes was dissected to expose the FT lumen and brushings were collected from the lumen epithelium. FTE cells were also cultured in NOSE-CM media. In parallel, during these procedures FT tissue was collected in formalin.

2.2.2 Growth Curves for FTE primary cell lines and X-gal staining for senescence

Growth curves for FTE cell lines were performed to calculate the population doubling times and also to determine the lifespan of a cell line before it senesces. 100,000 cells were plated in 60mm culture dishes in triplicate. The cells were regularly split before they reached confluency at a ratio of 1:3 to 1:7 and the number of cells was counted for each replicate at each passage. The population doublings (PD) were calculated using the formula: PD=log (total cell number/initial cell number)/log2. The cumulating population doublings were plotted and average population time was calculated based on the middle of the growth curve for each cell line. The lifespan of the cell lines was determined by the number of days in culture until they senesced.

When the cells stopped dividing and had vacuolated cytoplasms, senescence was confirmed by assaying for β-galactosidase expression using X-gal. The culture medium was aspirated and 2ml of 0.05% glutaraldehyde in PBS was added. The plate was then incubated for 5 minutes at room temperature and glutaraldehyde was discarded. Cells were washed with PBS and 2ml of PBS was added and incubated for 10 minutes at room temperature. PBS was aspirated and cells washed again in PBS. Cells were then covered with the X-Gal solution (1 mg/ml X-Gal in 5 mM $K_4Fe(CN)_6 \cdot 3H_2O$/5 mM $K_3Fe(CN)_6$/1 M $MgCl_2$/PBS, Promega) and incubated at 37°C, 5% CO_2 for 2-48 hours until the

cells developed a visible blue culture. The stained cells were visualized and photographed under a light microscope.

2.2.3 Karyotyping of FTE cell lines

FTE cells were plated at 40-60% confluency in a 25cm^2 flask. The karyotypic analysis was performed at The Doctors Laboratory (TDL). The cytogenetic profiles were analysed and commented by a senior Cytogeneticist (Mr. Terry Ballard).

2.2.4 Immunofluorescence cytochemistry of cultured cells

NOSE, FTE and control cell lines were grown on coverslips. When they reached ~80% confluency cell monolayers were washed with ice cold PBSAg (PBS with 0.3% fish skin gelatin, Sigma). Cells were then fixed for 10 minutes in cold 3% formaldehyde (VWR) prepared in PBS. The fixative was aspirated and fixed cells were washed with ice cold PBSAg. The cells were permeabilised by incubating for 5 minutes at room temperature in 0.3%v/v Triton X-100 (Sigma) in PBSAg and rinsed twice with PBSAg. In order to block unspecific binding of the antibodies each cell monolayer was incubated with 50µl 1% v/v non-immune goat serum (Invitrogen) for 10 minutes at room temperature. After aspiration of the blocking serum 50µl of optimally diluted primary antibody in PBSAg was applied to the coverslips. A summary of all the primary antibodies and their dilutions used for cell line characterisation by immunofluorescence staining is shown at table 2.3. The cells were incubated with primary antibody at room temperature for 1 hour. After rinsing the primary antibody with PBSAg, the appropriate Alexa Fluor-488 (1:2500 dilution, Invitrogen) secondary antibody (Table2.3) was applied and incubated for 20 minutes in the dark. After rinsing thoroughly with PBSAg, Evans Blue (Sigma) solution (Prepared with 1µl of the stock in 10ml of PBS) was applied to the coverslip for 10 minutes in order to stain the cytoplasm of the cells. The coverslips were then rinsed with PBS and transferred to labelled slides to add 5µl of DAPI (Vector Laboratories Inc) used for nuclear staining. Mounting medium (Sigma) was quickly added and the cells were covered with a coverslip. The slides were incubated for 30 minutes in

room temperature in the dark before visualising under an Olympus BX64 fluorescence microscope. If not observed immediately slides were kept at -4°C protected from light for maximum of one week.

Antibody	Dilution	Source
Anti-Ck7 (clone LP5K)	1:1000	CRUK
Anti-pan cytokeratin (clone AE1/AE3)	1:1000	Dako
Anti-BerEP4	1:1000	Dako
Anti-CA125	1:1000	Dako
Anti-E-Cadherin	1:2500	Cell Signalling
Anti-FSP	1:1000	Sigma
Anti-Factor VIII Related Antigen	1:1000	Neomarkers
Anti- PAX8	1:1000	Proteintech
Anti-Laminin	1:1000	Sigma
Anti-Vimentin	1:1000	Abcam
Secondary Anti-mouse IgG	1:2500	Invitrogen
Secondary Anti-rabbit IgG	1:2500	Invitrogen
Secondary Anti-mouse IgG	1:2500	Invitrogen

Table 2.3**:** Antibodies used to characterise NOSE, FTE and control cell lines

2.2.5 Immunohistochemistry of cultured cells

Cells cultured in 100mm dishes in the appropriate culture medium were trypsinised and counted as previously prescribed. Cells were seeded on 16 chamber slides (NUNC) at a 2×10^3 density and incubated at 37^oC, 5% CO_2 for 24-48 hours until they reached 60-80% confluency. The cell culture medium was then removed and replaced with neutral-buffered formalin (VWR). The samples were then processed into paraffin for immunohistochemistry at the University College Hospital Histology Services. Immunohistochemistry was performed at University College Hospital Advanced Diagnostic Laboratory.

Due to the high cost of the 16 chamber slides some cell lines were prepared in two dimensional agarose gel cell pellets for paraffin processing. Cells were cultured in 100mm dishes under normal culturing conditions until they reached 90% confluency. For the preparation of the pellets 2% agarose (Sigma) was dissolved in distilled water. Agarose was allowed to cool down enough to not kill the cells when added to the pellets. Cells were trypsinised, pelleted and the media removed. The pellet was then thoroughly flicked and one drop of 2% agarose was added with a Pasteur pipette directly on the pellet with constant agitation so that the cells were mixed and diluted in the gel. The

pellet was allowed to cool for 15min at room temperature, removed from the tube and fixed in formalin. The agarose pellet was then processed in paraffin and stained with appropriate markers as before.

2.2.6 Three-dimensional (3D) culturing of FTE cells and processing for immunohistochemistry

In order to generate 3D cultures, 1×10^6 FTE cells were cultured in 100mm tissue culture dishes previously coated with 2.5% Poly-hydroxymethylacrylate (polyHEMA) solution (Sigma) to prevent cells' attachment to the surface of the vessel. 2.5% polyHEMA solution was made in 95% ethanol. The solution was then heated at ~65ºC until fully dissolved (1-2 hours). Under sterile conditions the plates were coated with the PolyHEMA solution and allowed to evaporate. A second coat was then applied to ensure complete covering of the surface. Before use the plates were washed with PBS twice. The cell line to be cultured in 3D was first cultured in 2D at <90% confluency. Cells were trypsinised and counted. The cells were thoroughly resuspended to ensure single cell suspension and between $1x10^6$ cells were cultured in 100mm double coated dishes. The cells were then incubated for 15, 26 or 40 days at 37^oC, 5% CO_2. The media was carefully replaced every 3-4 days by tilting the plate, letting the spheroids settle and carefully removing up to 80% of the culturing media before slowly replenishing with fresh.

The 3D cultures were then fixed and processed for immunohistochemistry as follows. The 3D cultures were removed from the incubator and culture medium containing the spheroids was poured into 50ml v-bottomed tube. Spheroids were left at room temperature to settle for 20 minutes. The cell culture medium was then removed and replaced with neutral-buffered formalin (VWR). The samples were then processed into paraffin for immunohistochemistry at the University College Hospital Histology Services similarly to FT tissue previously collected. Immunohistochemistry of the FTE 3D cultures and FT tissue was performed at University College Hospital Advanced Diagnostic Laboratory. Representative slides for each cell line were also stained with Heamatoxylin and Eosin (H&E) to visualise the architecture of the 3D cultured spheroids.

2.3 Investigating loss of heterozygosity (LOH) & allele specific LOH

2.3.1 Microcell-mediated chromosome transfer of chromosome 18

The experimental procedure of microcell-mediated chromosome transfer of chromosome 18 (MMCT-18) was performed by Dimitra Dafou and is extensively described in Dafou et al, 2008. Briefly, human monochromosome hybrids were generated in mouse A9 cells. The A9:monochromosome donor hybrids were exposed to colcemid for 48 hours to induce micronucleation. The hybrid donor cell lines, tagged with Hytk selectable marker, were fused with EEOC cell lines TOV112D and TOV21G (American Tissue Culture Collection). TOV21G and TOV112D hybrid clones were selected with hygromycin B (Sigma). The tumourogenicity of the resulting hybrid clones was assessed with anchorage independent growth in soft agar and matrigel invasion assays.

The parental cell lines and two clones of each cell lines that showed reduction of neoplastic phenotype were used in gene expression microarray analysis, microsatellite genotyping, cytogenetic analysis to confirm uptake of complete or partial chromosome 18 in MMCT hybrids. 29,098 genes were screened by 32,878 probes in the microarray analysis. An Anova test was used to generate P values for differential gene expression between the hybrids and parental. These were further adjusted for multiple testing. The results were ranked to finally lead to generation of a list of candidate genes significantly up or down regulated in the hybrids. Nine candidate genes were selected to be used for genotyping tagging SNPs within them and genotyped in the iPLEX genotyping platform.

2.3.2 Malova collection of samples

MALOVA (Malignant Ovarian Cancer prediction) is a Danish collection of samples and it consists of 446 women of 30-80 years of age that have been diagnosed with EOC between 1994 and 1999. All the cases were recruited into the study at surgery before diagnosis of the disease, therefore they are said to be incident cases. Follow-up has been done of the patients to collect survival

information and until now 201 deaths have been confirmed. Tumour samples of all cases were paraffin embedded and made in tissue array slides. Genomic DNA from the cases was extracted from pre-operative blood samples by Whatman International Ltd with chloroform protocol (Ely, UK). 301 tumours and their matching germline DNA were used for the LOH study.

2.3.3 DNA extraction from Malova tissue array slides

Methylene green solution (Sigma) was prepared by adding 1-1.5ml of methyl green in 200ml of distilled water. The Malova tissue array slides were then used for methyl green staining. Methyl green staining of each slide was then performed as follows: Xylene (Sigma) for 5 minutes, then again in Xylene for 5 minutes, then in 100% ethanol for 5 minutes, then in 85% ethanol for 5 minutes, then in 75% ethanol for 5 minutes, then rinsed in distilled water, then in methyl green solution for 30sec to 1minute until tissue was slightly green, then in distilled water for 5 minutes twice. The slides were then left to air dry for 15-30 minutes until completely dry.

Extraction Buffer (EB) was prepared for all extractions to be made in the same batch. 30µl EB was prepared for each sample of the tissue array slides to be micro-dissected. The EB was prepared on ice and kept on ice while being used. EB buffer recipe is shown below. Some of the reagents used for the EB (MgCl2 and 10×Buffer) were from the Amplitaq Gold kit (Applied Biosystems).

Reagent	**Volume (µl)**
dH2O (distilled	25.54
10xBuffer	3
MgCl2	1.8
Proteinase	0.48
Tween-20	0.15

The methyl green slide was then placed on the microscope. The marked area to be dissected was located, and a sterile 21G 11/2" needle, 0.8 × 40 mm, (BD Microlance) was dipped into EB and the tissue scraped (one 10x field is required) trying to avoid any stroma or necrotic tissue. For best results the needle was held with the curve facing against the direction of scraping. The flakes of tissue were rinsed in the EB, until the whole area was extracted. Each

sample was vortexed for 5-10sec and then spun down. The samples were then incubated at 55^{o}C for 24 hours. After this incubation the samples were spun down, and 1µl Proteinase K (Promega) was added in each sample and then vortexed, spun down and incubated for further 24 hours at 55^{o}C. Samples were then vortexed and spun down again. Proteinase K was inactivated by heating the paraffin samples at 99°C for 15 minutes. Then the samples were placed on ice for 10 min and were either used straight away or stored at -20^{o}C.

2.3.4 Testing the efficiency of DNA extraction from Malova tissue array slides

The efficiency of the DNA extraction following micro-dissection of the tissue array slides was assayed by performing a PCR against BRCA1 exon 17. The following primers were used:

Forward 1.17F agctgtgtgctagaggtaactc

Reverse 1.17R gtggttttatgcagcagatg

The following PCR mix was prepared and 2µl of DNA was added for each reaction.

Reagent	**Volume (µl)**
dH_2O	6.38
10X PCR Buffer (Applied Biosystems)	1.5
25mM MgCl (Applied Biosystems)	1.5
10mM dNTPs (Invitrogen)	1.5
BRCA1.17 _F primer	1
BRCA1.17 _R primer	1
AmpliTaq Gold (Applied Biosystems)	0.12
Total volume before DNA addition	13

The PCR program was then performed at 55^{o}C, 40 cycles. Then the PCR product was subjected to electrophoresis on a 1.5% agarose gel.

2.3.5 iPLEX Genotyping assay (Sequenom) for tSNPs from MMCT-18 genes

Haploview software was used to generate tagging SNPs (tSNPs) across the nine candidate genes by Lydia Quaye in context with an association study she was performing at the time. The Sequenom MassARAY iPLEX multiplex genotyping platform was used to genotype the genomic and matching tumour Malova extracted DNA samples for the MMCT-18 LOH study. A list of the primers used for this experiment is shown in Table 2.4. 120µl of 500nM of the forward and reverse primers of all the SNPs were combined in a primer mix for a final concentration of 100nM in each reaction of 5µl. Additionally a dNTP mix was prepared with 100nM of each of the dATP, dCTP, dGTP and dTTP. Then a PCR mastermix was prepared for each sample: (1.85µl distilled Milli-Q water, 0.625µl of PCR buffer with 10X magnesium chloride [$MgCl_2$], 0.325µl of 25mM $MgCl_2$, 0.1µl of 25mM dNTP mix, 1µl of primer mix [500nM of each primer], and 0.1µl of 5U/µl Hotstar Taq® DNA polymerase enzyme), and added to 10ng of germline DNA or dried 2µl of the tissue array extracted tumour DNA. The PCR program used was the following: 94^{o}C for 15 minutes, cycled 45 times (at 94^{o}C for 20 seconds, 56^{o}C for 30 seconds, 72^{o}C for 60 seconds), and inactivated at 72^{o}C for 60 seconds.

The product of this PCR pre-amplification mixture was then treated with a shrimp alkaline phosphatase step (SAP) to dephosphorylate the unicorporated dNTPs in order to prevent them from being incorporated in following primer extension reactions that would lead in having results with contamination peaks.

This cleaning step was performed by incubating the PCR products with 2µl of SAP mix (1.53µl distilled Milli-Q water, 0.17µl of 10x SAP buffer and 0.3µl of 1U/µl SAP enzyme for each sample). The incubation steps were 37^{o}C for 20 minutes and 85^{o}C for 5 minutes. The dephosphorylated mixture was then cooled to 4^{o}C.

The extend primers were then pooled into four groups according the mass of the extend primers. The signal-to-noise ratios of the extend primers decrease with increasing extend primer mass, therefore these adjustments in extend primer concentrations were required in order to equilibrate the signal-to-noise ratios of the extend primers of different masses. The extend primers were

arranged into increasing masses, and the primers were split into 4 groups. Therefore, lower mass primers were grouped with other low mass primers and high mass primers were grouped with other high mass primers. The final concentration of the lowest mass primers was half of those in the highest mass group. Thus an extend primer mix was prepared whereby the final concentrations of the group 1 extend primers (lowest mass) was 0.625μM, group 2 was 0.833μM, group 3 was 1.042μM, and group 4 (highest mass) was 1.25μM.

An iPLEX reaction mix was prepared (for each sample): 0.755μl distilled water, 0.2μl of 10X iPLEX buffer, 0.2μl of iPLEX termination mix, 0.804μl of the extend primer mix and 0.041μl of the iPLEX enzyme. 2μl of this cocktail was added to each sample. The mixture was mixed and covered with adhesive seal. This was subsequently cycled for the following PCR conditions:

HOLD	40 CYCLES			HOLD	HOLD
	Hold	5 cycles			
94°C 30 seconds	94°C 5 seconds	52°C 5 seconds	80°C 5 seconds	72°C 3 mins	4°C 15 mins

The iPLEX reaction products were desalted by adding 25μl of water and 6mg of Clean resin (using a dimple plate). A nano-dispenser was used to dispense the iPLEX reaction products onto a 384-element SpectroCHIP bioarray. The SpectroCHIPs were read on Autoplex. The genotypes were called using the Typeranalyzer software and the peak heights exported to excel for further analysis. The peak heights were used to analyse LOH and allele specific LOH.

2.3.6 Analysis for Loss of Heterozygosity (LOH)

The ratio of the allele peak heights between the tumour and the germline DNA for heterozygous (informative) individuals was used to determine LOH (Canzian et al, 1996). The formula L= (at2 x an1)/(at1 x an2) was used, where at1,at2 and an1,an2 are the peak heights of the 2 alleles of the tumour and the germline DNA respectively. A value of $L<0.6$ and $L>1.67$, would indicate that one of the alleles has decreased more than 40% resulting in LOH (Canzian et al, 1996). The calculated frequency of overall LOH for a specific gene was based on the combined analysis of multiple tSNPs within the gene. LOH was

recorded if any informative SNP in a gene showed LOH, even if other informative tSNPs did not show LOH. Allele specific LOH (AS LOH) was assayed based on the deviation from the expected 1:1 random ratio of loss of one allele compared to the other allele. Evidence of AS LOH would be shown by preferential loss of one allele compared to the other. The P value was calculated using a two-sided Fisher's test. An example of an output form the iPLEX genotyping result where the peak heights for the two alleles are demonstrating LOH between the tumour and germline DNA of a sample for SNP rs144848 in BRCA1 is presented in Figure 2.1.

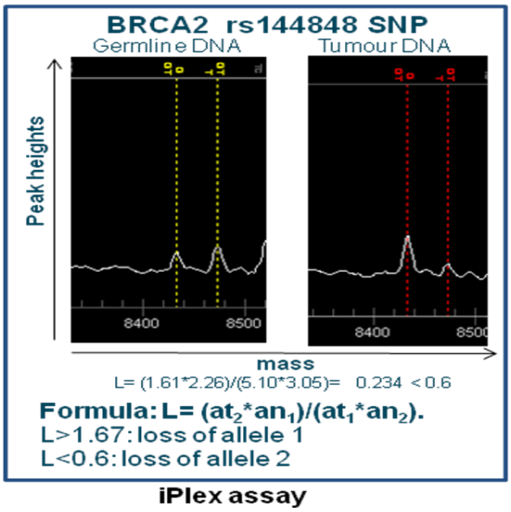


Figure 2.1: Example of LOH based on allele peak heights generated by the iPLEX genotyping platform. Presented is an iPLEX histogram where rs144848 of BRCA2 was genotyped in the same patient's tumour and germline DNA/ The peak heights of the 2 alleles are used in the formula to determine whether there is loss of heterozygosity.

2.3.7 Survival analysis

Survival analysis was performed to establish whether LOH status was associated with Malova patients' survival. Similarly to the analysis of overall LOH across genes, LOH was recorded if any informative (heterozygous) tSNP in a gene showed LOH, even if other informative tSNPs showed no LOH.

Univariate analysis: Survival analysis relatively to LOH was performed in STATA statistical software using Cox proportional hazards regression model to measure the effect of LOH over the time death takes to occur within the specified 10 year period and Kaplan Maier curves were generated. The STATA commands I used for survival analysis were: i) **set mem 290m**, ii) **stset yearoutcome, failure(outcome) enter(yearenter) exit(time 10)**, iii) **stsgraph, by(lohstatus)**, iv) **xi:stcox lohcode**

Multivariate analysis: Cox regression analysis in STATA was also used to estimate the effects of age at diagnosis, histological subtype, stage and grade of the tumour on the 286 samples used for the analysis, using the <40 age group, serous histology, grade 1 and localised/early stage as calibrators/references to compare the rest with for each prognostic factor respectively. The analysis was performed in STATA using the command **xi: stcox i.histology i.agegroup i.grade i.stagegroup.** The results of all genes were adjusted for the prognostic factors age at diagnosis >60 and advanced stage that were both significantly associated with worse survival. This kind of analysis is termed as multivariate analysis. Cox regression multivariate analysis was performed to assess the effect of LOH of each gene on survival in all subtypes, serous only and non-serous only subtypes adjusted for prognostic factors where appropriate. The commands used in STATA software where: i) **rename _Iagegroup_4 age4** and/or **rename _Istagegro~2 stage2**, ii) **xi: stcox GENElohstatus age4 stage2.**

Plex	SNP	Forward primer	Reverse primer	Extend primer
28plex	rs2280201-PCR1	ACGTTGGATGTCTGCATTATCTTAGGAAC	ACGTTGGATGCACACTCCCAGTTTATGTGA	TCTGCCCTGAGCAAA
28plex	rs11655966-PCR1	ACGTTGGATGTCCCTCCTAATCAGTGTGCC	ACGTTGGATGCTTTACCTGCAATTGAGGGG	GAGCAGGGAGCAAAA
28plex	rs2394655-PCR1	ACGTTGGATGCCCACACGTAGGTCCTTCTA	ACGTTGGATGAGCGAGGTGGGAATCCCAGT	AATCCCAGTGTGTGAG
28plex	rs7239066-PCR1	ACGTTGGATGAGAGAAGGGGATGGATTTAC	ACGTTGGATGCATCCCTAACTTCCATATCC	CCATATCCAAGCCATCA
28plex	rs11082221-PCR1	ACGTTGGATGCCTTCCCTCAATTCTGATGC	ACGTTGGATGGGGAACCCCAATTTAGAGTC	AGAGTCTGTCCAAAGCA
28plex	rs3750772-PCR1	ACGTTGGATGCCAGGTTCTAGTGCCATCC	ACGTTGGATGTCTAGGCACCCGAGGCAGC	TCCACCACCACACTGGTT
28plex	rs2246713-PCR1	ACGTTGGATGGGTGGGTGGATAGAGATGTG	ACGTTGGATGGAGGTAGAACCAAAAGCCTG	AAAAGCCTGAAGGGGCTA
28plex	rs7189819-PCR1	ACGTTGGATGATTCCTTTACATCACCCAGC	ACGTTGGATGCTAGTTTTAGCTCTGCCACG	GCCCCGCACTGTCATTCGC
28plex	rs4541111-PCR1	ACGTTGGATGAGGCACGTTCAGCAGGCTA	ACGTTGGATGATGGGTTTGCGGTGCACAG	CCATTGCACAGCGCCTGGT
28plex	rs975590-PCR1	ACGTTGGATGACAGGGAAGGTGCCATGAGT	ACGTTGGATGTGAGGACTCGTTTCCTTGTG	AGGTATGAGGCTGTCCTGG
28plex	rs3923086-PCR1	ACGTTGGATGCCAAAAAAGCCAAAACAGCC	ACGTTGGATGTACCACCAAGATAGCAGGAC	GCATGGAAGAAAGAGTGTT
28plex	rs9821568-PCR1	ACGTTGGATGGCTATGGATCCTTTCTCTGG	ACGTTGGATGGCCCAAATACCTCTTGAAGT	TACCTCTTGAAGTTGGGAAC
28plex	rs2894111-PCR1	ACGTTGGATGTGGCATCAGCAACCTTGCCT	ACGTTGGATGCACCCTCCCACCTAATGTAT	GGCCACCTAATGTATTCCCCC
28plex	rs9931702-PCR1	ACGTTGGATGCAACTTATGAAATGGCTAAC	ACGTTGGATGGTATGTAATGAGCAGTCTCC	TGAGCAGTCTCCAAATGAAGA
28plex	rs793477-PCR1	ACGTTGGATGCTGTGTCTCTTTGAGCAAAC	ACGTTGGATGCAATCCCCAGATGAAGATTAC	CAGAGGAAACTGAGGTTTACA
28plex	rs3181328-PCR1	ACGTTGGATGGCTGAACATCTAACAGTTTTG	ACGTTGGATGGGTCTAACATTGTTTCCCCT	AACCAAACATCACAATCCTCTG
28plex	rs2394656-PCR1	ACGTTGGATGTGGAGAGCAAGGAGGCACAG	ACGTTGGATGAGATCAACAGCTCCGCCTAC	GCATCTCCGCCTACCGCAAAGC
28plex	rs7650365-PCR1	ACGTTGGATGGGGGTGTGTTGGTAACTTTTG	ACGTTGGATGCCAACTAGCCTATGACCTTG	GATCCTATGACCTTGATCATCTA
28plex	rs6788750-PCR1	ACGTTGGATGACTCTTCCACAATATGAAGG	ACGTTGGATGGATAATCTTTGGACTCCTTG	CTTGGGACTCCTTGAAGGTAGTT
28plex	rs10999152-PCR1	ACGTTGGATGTGTCTTTTCCCAGACTGTCG	ACGTTGGATGTAGGCGTTGGAGGGATAACT	GAGGGATAACTTGGAGAGAAAAC
28plex	rs17446518-PCR1	ACGTTGGATGCATGGATTGAATGAGCACTG	ACGTTGGATGAGATGCTTTCAGCAGGCAAG	GCTGTCAGCAGGCAAGGGTTCATA
28plex	rs13063604-PCR1	ACGTTGGATGAATCCAGGCTTTCTGATCCG	ACGTTGGATGAGAGATTTCTAGCAGGAGGG	GGGCGGAGGGAAGCTGGGTGGCAC
28plex	rs9651713-PCR1	ACGTTGGATGCCTCCTAGTTCACACCACAT	ACGTTGGATGACCAGGATGGTACAGGATTC	GTACAGGATTCAATATTCTATCAAA
28plex	rs7908957-PCR1	ACGTTGGATGTCCTCTGGGATGTGCATGTG	ACGTTGGATGCCCTGCTCTGAGAGCTCTAA	TTTCGCTGAGAGCTCTAAATGGCAC
28plex	rs11868547-PCR1	ACGTTGGATGGGAGCTCTGATGCCTTGAAC	ACGTTGGATGGGTATGAACCCCCATTCCAG	GGGCGGGCACAGCACCTGCATGGTTA
28plex	rs144848-PCR1	ACGTTGGATGATCTGAAGTGGAACCAAATG	ACGTTGGATGGTCACTTCCACTCTCAAAGG	CCCTTCCACTCTCAAAGGGCTTCTGAT
28plex	rs799917-PCR1	ACGTTGGATGAAGGTTTCAAAGCGCCAGTC	ACGTTGGATGAGAGTGGGCAGAGAATGTTG	CCGGCTGCATTTCCTGGATTTGAAAAC
28plex	rs17801966-PCR1	ACGTTGGATGGGTTCTACGGTTAGAACAGC	ACGTTGGATGATTAGTGCCCCTCAAGTCTG	CAAGTCTGATTTAATGATATTAGACTAG
33plex	rs9864437-PCR1	ACGTTGGATGAGCTGGTGTCCTTTCTGCG	ACGTTGGATGTTCCCTTCTCTCATCCACCC	CCCTTCCCCAGTGAC
33plex	rs4295944-PCR1	ACGTTGGATGAGCTCATCTGGGACCCTCA	ACGTTGGATGAAGGGGTTGTGCCCCAGAC	CCAGACCCGCAAAGA
33plex	rs2271695-PCR1	ACGTTGGATGCGCTGCTTCACAGAGTTGAC	ACGTTGGATGCAAGATGGCCTATCTTGCCG	CCCTCCACGCCAACAT
33plex	rs10999147-PCR1	ACGTTGGATGATGCCTTCAGTGCACAGCTC	ACGTTGGATGTAATGAGGTTTCCAGCCAGC	CAGGCCTATGAGGACA
33plex	rs1637001-PCR1	ACGTTGGATGAGAGTTTTGGTACCAGAGGC	ACGTTGGATGTTTAGAAAATCCTGTCTCCC	CCTGTCTCCCATTTTCC
33plex	rs11762932-PCR1	ACGTTGGATGCTCTGCATGTGGTTATTCGG	ACGTTGGATGAGTGCAATAAAGCGAGATAC	AAGCGAGATACACCTGT
33plex	rs4857836-PCR1	ACGTTGGATGTTTTGCCCTTGCAGGGGACT	ACGTTGGATGATGCCAGTAGCACCCCCAA	TCTCACCCCCAACCAAGT
33plex	rs7591-PCR1	ACGTTGGATGACCTTTCCTCCACACACCTT	ACGTTGGATGATGGCAAACAGAATGTACAG	TAACACAAAAACCCGAAC
33plex	rs3743772-PCR1	ACGTTGGATGTTGTTTGTCTCTGCCCCATC	ACGTTGGATGTTACGTAACAGCGCCTGGGA	TGTCAGTCATGGAAGTAT
33plex	rs1053495-PCR1	ACGTTGGATGCCCACATAGAAGCCACTGAT	ACGTTGGATGTTGCAGGTGCACTGACGTTC	GGCCATGGGGAGAAATGA
33plex	rs17338680-PCR1	ACGTTGGATGCCTTTTAACTACTCTCCATC	ACGTTGGATGCCTATTCTATGGGTGAAGGC	GAAGGCTGCTAGGGTATTG
33plex	rs11079571-PCR1	ACGTTGGATGTGGAGGGGTTGTGACTCAG	ACGTTGGATGAGGGACCAGAAGTCAGAGGC	CCTTCTTGGCTGCAGTGCTC
33plex	rs11618371-PCR1	ACGTTGGATGGCCGATATTGTAAGGCCCTC	ACGTTGGATGCGAAGGACCTAGAACAGTGA	CTCGAATTTGGGGAACCCTG
33plex	rs7210356-PCR1	ACGTTGGATGGTTCACACCTGGGAATAGAC	ACGTTGGATGCAAGACAGAGTTCCAACTCC	TCCCCCTCCTCAGCTGTGTGC
33plex	rs2282657-PCR1	ACGTTGGATGGGAGTTGAATATATTCTGG	ACGTTGGATGCCAAAAGATCACCAGTGTAA	TCACCAGTGTAAAACGTAAGG
33plex	rs793446-PCR1	ACGTTGGATGGCAAGGCACTGTACTTGCTA	ACGTTGGATGATGATCAGACTGATCTGTGC	AGGTTTTGGGGAAAAAAGAAC
33plex	rs3783194-PCR1	ACGTTGGATGATTTTCTCCCTAATAGCAC	ACGTTGGATGGGGTGATGAGTGGGATTTAG	CTAGTGGGATTTAGTTTCTCTA
33plex	rs523104-PCR1	ACGTTGGATGGTCAGAGACTCTCCAGCATC	ACGTTGGATGTCTTGCAAACAGAATCTGCC	GGGGATGACTGTGAAGAGATGA
33plex	rs3783197-PCR1	ACGTTGGATGGGAAGACTTTTGGCTTGTTG	ACGTTGGATGTGATGGCCTGTCAGGCTTTT	CTCAATGAACTGATTAACCACAA
33plex	rs2394644-PCR1	ACGTTGGATGGTGTATGTGTGTTGTGGCAG	ACGTTGGATGTAAGAGAGCGGAACAAACCC	GGGCGCGGAACAAACCCTGATTG
33plex	rs13091198-PCR1	ACGTTGGATGCTTTCACTTAGCATGATGCTG	ACGTTGGATGTAGAATACTATTCAGCTGG	GCTGGAAAAAGAAATAAGTACTA
33plex	rs12494994-PCR1	ACGTTGGATGGCATTGCACTGTATTAGAAGC	ACGTTGGATGTCTTGTTTGAATTTCCCAGG	CCCTTGCAGTTTGTAATAGCTGCT
33plex	rs3181174-PCR1	ACGTTGGATGCGATCATGAACATAAAGCAAC	ACGTTGGATGAACTTTCCCCACAGAGAAGG	AGGAGGGGAGTTTATTTTGTATAG
33plex	rs3732402-PCR1	ACGTTGGATGGGTCACCAAAAGGCTATCAC	ACGTTGGATGGATGATAGCCATTAGGGAGG	CTCTCGCCATTAGGGAGGCTACCAG
33plex	rs3181175-PCR1	ACGTTGGATGACAAGATTTCTTTTCATG	ACGTTGGATGTATGTCTTGGTTGGGTCACG	GTCACGTAACCAAACATAAAAGTAT
33plex	rs9860614-PCR1	ACGTTGGATGCCGTAAGTCCATTCTGATAC	ACGTTGGATGTGCTCACCTCTAGAGAACTG	GAGAACTGAGTATTTACCAAAATTAT
33plex	rs4791171-PCR1	ACGTTGGATGAGCCACCACAACCTGCCAAA	ACGTTGGATGCAGTGCTGCAAGGATTTTGG	GTGATGCAAGGATTTTGGTGGGAGTC
33plex	rs9304261-PCR1	ACGTTGGATGGCTCTTCTCTAGGTCATATC	ACGTTGGATGCTATGAAAATACAGTCTTTCT	TTTGTTTTTTTAAAAAAAGACTTAAAT
33plex	rs6480440-PCR1	ACGTTGGATGACGGCTGGCACTCATAAGAG	ACGTTGGATGCTGTTTCCCAAGAAGTTCTC	GGCGCAGAAAAGCTGTTTAATGGAAGA
33plex	rs518604-PCR1	ACGTTGGATGATCCCAGTACTGGCTGTCTT	ACGTTGGATGCTCTTTGTAGGATCAAGTGG	CATGGAGAAGAGGGCTTGAAGTTGATC
33plex	rs995845-PCR1	ACGTTGGATGGATAACCAACAATTTGTCTTC	ACGTTGGATGTGCCTGGCTGTTTTCTTTAC	CCTTACTGGCTGTTTTCTTTACTATAGA
33plex	rs796977-PCR1	ACGTTGGATGTGGCAATCACGGTGTAGTAG	ACGTTGGATGTTCTCTGGCAATCTCATGGC	GGATATTAATTTGGGACACTACTCTTGT

Table 2.4: Summary of the primer sequences used for iPLEX genotyping of the tSNPS in the MMCT-18 candidate genes.

2.4 Evaluating Gene expression

2.4.1 Cell line RNA extraction

The RNeasy Mini Kit was used for RNA extraction of EOC, NOSE and FTE cell lines according to the manufacturer's instructions. A NanoDrop® Spectrophotometer was used to determine the concentration of the extracted RNA by reading the absorbance at 260nm. The ratio of absorbance at 260nm and absorbance at 280nm was used for purity estimation.

2.4.2 Reverse transcription of cell line RNA into cDNA

Quantified RNA samples were normalised to 100ng/µl in RNase free H_2O in 0.5ml eppendorf tubes. 1µg of RNA was used for the reverse transcription reaction. 1µg of random primers (Promega) and 15Units (0.35ul) of RNasin (Promega) were added and the total volume of the reaction brought to 14µl with RNase free H_2O. This step introduces first strand synthesis. The samples were then heated at 75^{o}C for 5 minutes in order for secondary structures to melt and immediately cooled to prevent reformation of the secondary structures. The samples were then briefly spun down. 5µl of M-MLV Reverse Transcriptase reaction buffer (Promega), 1µl of M-MLV Reverse Transcriptase enzyme and 5µl of 2.5mM dNTPs (Promega) were added to the sample which were then be incubated for 10 min at room temperature. The reactions were incubated at 45^{o}C for 50 minutes**.** The cDNA was then diluted down to 5ng/µl by addition of 175µl RNase free H_2O (Promega). The efficiency of the reverse transcription (RT) reaction was then assayed by β-actin PCR. Human female DNA, (Promega) was used as a positive control for the PCR reaction. The primers used for β-actin used were the following:

Forward β-actin_F: GTCCTCTCCCAAGTCCACAC
Reverse β-actin_R: GGGAGACCAAAAGCCTTCAT

10ng of cDNA was added in a 96 well PCR plate (2µl of the 5ng/µl stocks) containing the following reaction mix. The following recipe was used to prepare the PCR reactions:

Reagent	Volume (µl)
H_2O	7.38
10X PCR Buffer	1.5
25mM MgCl	1.5
10mM dNTPs (Promega)	1.5
β-actin_F	0.5
β-actin_R	0.5
AmpliTaq Gold (Applied Biosystems)	0.12
Total volume before DNA addition	13

A standard PCR program was performed for 60^{o}C annealing temperature, 40 cycles. Finally, 8µl PCR product +2ul loading dye is run on a 1.5% agarose gel to confirmed production of cDNA during reverse transcription reaction.

2.4.3 Optimisation of Taqman® Real time PCR expression assay to evaluate differential expression of candidate genes

It is known that different endogenous controls are more appropriate to use for the normalisation of expression assays in different tissues, therefore a panel of 4 endogenous controls (housekeeping genes) were initially tested in NOSE and EOC cDNA samples. These were β-actin, GAPDH, TBP (TATA binding protein) and 18S (18S ribosomal RNA). GAPDH and β-actin were chosen as the more reliable for these samples as they exhibited no significant variation between samples. The expression of the candidate genes was normalized against those two endogenous controls.

The next step in the optimisation was to check whether multiplexing of the endogenous controls with the target probes would provide reliable results. The standard curves were more reliable when the probes were not multiplexed therefore I decided to perform downstream experiments using “2 tube reactions” where the endogenous and the target probe were assayed in separate tubes.
Other optimisation experiments were also performed. A Real time experiment using cDNA produced with oligodTs versus random hexamers was tested. No difference was found between the two different cDNA preparations. Assays were also optimised with respect to the minimum amount of probe needed. The optimised reactions recommended by Applied Biosystems were 20µl reactions. For the purpose of my experiments I initially scaled down the reaction to 10µl

and subsequently to 5µl reactions after drying the cDNA. Two probes were used to test that drying the cDNA would provide reliable relative expression data. The curves produced suggested reliable consistent data.

2.4.4 Evaluating gene expression by performing Real time PCR in ABI7900 Taqman® Sequence detection system

4ng of EOC and NOSE cDNA (4ng) were plated in 384 well PCR plates (VWR) and then dried. All samples were run in triplicates. Housekeeping genes were used as loading controls. Additionally, 8 serial dilutions of a mix of the cDNAs used starting from 5x more concentrated cDNA than the samples to be run was prepared in order to be used to create a standard curve that would confirm the quantity of cDNA versus expression linear relationship.

Taqman® Real time PCR probes and Taqman® Real time PCR Master Mix (Applied Biosystems) were used for the reactions. Taqman® gene expression probes were obtained from Applied Biosystems for *SKAP1* (hs00175372_m1), *TIPARP* (hs00604498_m1), *HOXD3* (hs00232506_m1), *BNC2* (hs00417700_m1) ,*HOXD1* (hs00707081_s1), *MYC* (hs00153408_m1), *MERIT40* (hs01572961_m1), *ANKRD41* (hs00401326_m1), STAG3 (hs00429370_m1). All assays had a FAM^{TM} reported dye at the 5' end of the probe and a non-fluorescent quencher at the 3' end of the probe. The mastermix used contained buffer, dNTPs, passive reference dye, thermostable hot-start DNA polymerase, and other components formulated for reliable TaqMan® Assay-based real-time PCR. The assays were carefully chosen to make sure that they would detect the maximum number of transcripts.

The following reaction mix was prepared:

Reagents	**Volume (µl)**
dH_2O	2.25
Taqman® probe (20×)	0.25
Taqman® Universal PCR Mastermix (2x)	2.5

The Real time PCR expression assays were performed using the Applied Biosystems7900 Taqman® Sequence Detection System and the following

program ([50^{o}C (2min), 95^{o}C (10 minutes), 50 cycles: 95^{o}C (15sec) & 60^{o}C (1min)].

2.4.5 Evaluating gene expression by performing Real time PCR in a Fluidigm 96.96 Dynamic array.

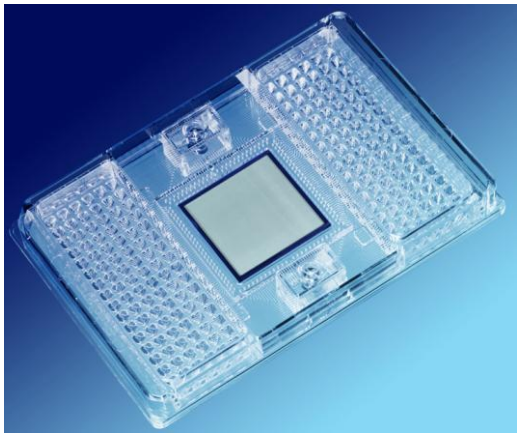
A large number of gene expression assays was multiplexed using Taqman® Real time PCR probes in the Fluidigm 96.96 Dynamic array platform. A comprehensive list of all the Taqman® Real time PCR probes used for my study in this array is shown in Table 2.5. The selected assays were multiplexed in 7 different sets (Fluidigm chips). The experimental procedure involved one pre-amplification step and before the Real time PCR step. This experiment was run by Miss Eva Wozniak, the research assistant in our laboratory, and it included assays and samples provided from several different projects within our group including the ones for my study.

Protocol for pre-amplification

The extracted cDNA from the NOSE, FTE and EOC cell lines was used to perform a pre-amplification step, which is a technique that causes amplification of just the target genes to be measured quantitatively. The step of pre-amplification included a small dilution of primers for each gene and only performed 10 PCR cycles so that entering the plateau phase is avoided. However, pre-amplified cDNA provided with a sample containing increased number of desired transcripts in the range of 500 fold.

Protocol for Fluidigm 96.96 Real-Time PCR

The first step involved priming of the Dynamic Array IFC. The control line fluid was injected into each accumulator on the chip (picture of the chip is shown on the right). The chip was then placed into the Integrated Fluidic Circuit (IFC) controller and then



the Prime (136x) script was run to prime the control line fluid into the chip. The second step involved preparing 10x assays using the volumes indicated below:

Reagent	**Volume per Inlet (µl)**
20x Taqman® Gene Expression Assay (AB)	2.5
2x Assay Loading Reagent (Fluodigm)	2.5
Total Volume	5
Final Concentration (at 10x)	Primers: 9µM; Probe:2µM

The third step involved the preparation of the Sample Pre-Mix and Samples combing the reagents as follows:

Reagent	**Volume per Inlet (µl)**
Taqman®Universal PCR Mastermix 2x (AB)	2.5
20x GE Sample Loading Reagent (Fluodigm)	0.25
cDNA	2.25
Total Volume	5

The Taqman® Universal PCR Master Mix was then combined with the GE Sample Loading Reagent in a 1.5ml sterile tube-enough volume to fill an entire chip. A volume of 2.75µl of this Pre-Sample Mix was then aliquoted for each sample. In each of those aliquots 2.25µl (4ng) of cDNA was added making a total volume of 5µl. The fourth step of the assay was to load the chip. When the Prime (136x) script described at the first step was finished, the primed chip was removed from the IFC controller and 5µl of each assay and each samples was pipetted into their respective inlets on the chip. The chip was then returned to the IFC controller and the Load Mix (136x) script was run for the samples and the assays to be loaded into the chip. When this was completed the loaded chip was removed from the IFC controller and the blue protective film was peeled from the underside of the chip and the chip was ready to be run immediately according the Fluodigm Data collection Software instructions.

Gene	Study	Taqman® assay (AB)
KIAA1715	Post GWAS analysis	Hs00327995_m1
EVX2	Post GWAS analysis	Hs01372231_m1
HOXD13	Post GWAS analysis	Hs00968515_m1
HOXD12	Post GWAS analysis	Hs00706957_s1
HOXD11	Post GWAS analysis	Hs00360798_m1
HOXD10	Post GWAS analysis	Hs00157974_m1
HOXD9	Post GWAS analysis	Hs00610725_g1
HOXD8	Post GWAS analysis	Hs00251905_m1
HOXD4	Post GWAS analysis	Hs00429605_m1
HOXD3	Post GWAS analysis	Hs00232506_m1
HOXD1	Post GWAS analysis	Hs00707081_s1
MTX2	Post GWAS analysis	Hs00198047_m1
KCNAB1	Post GWAS analysis	Hs00963155_m1
SSR3	Post GWAS analysis	Hs00606481_m1
TIPARP	Post GWAS analysis	Hs00296054_m1
LOC730091	Post GWAS analysis	Hs00327629_m1
PA2G4P4	Post GWAS analysis	Hs01689228_gH
LEKR1	Post GWAS analysis	Hs00416272_m1
MYC	Post GWAS analysis	Hs00905030_m1
PVT1	Post GWAS analysis	Hs01069044_m1
BNC2	Post GWAS analysis	Hs00417700_m1
CNTLN	Post GWAS analysis	Hs01124982_m1
SP6	Post GWAS analysis	Hs01941471_s1
SP2	Post GWAS analysis	Hs00370726_m1
PNPO	Post GWAS analysis	Hs00216680_m1
ATAD4	Post GWAS analysis	Hs00256987_m1
CDK5RAP3	Post GWAS analysis	Hs00602528_g1
COPZ2	Post GWAS analysis	Hs00212698_m1
NFE2L1	Post GWAS analysis	Hs00231457_m1
CBX1	Post GWAS analysis	Hs01080635_g1
SNX11	Post GWAS analysis	Hs00203367_m1
SKAP1	Post GWAS analysis	Hs00175372_m1
HOXB1	Post GWAS analysis	Hs00157973_m1
HOXB2	Post GWAS analysis	Hs00609873_g1
HOXB3	Post GWAS analysis	Hs00231127_m1
HOXB4	Post GWAS analysis	Hs00256884_m1
HOXB5	Post GWAS analysis	Hs00357820_m1
HOXB6	Post GWAS analysis	Hs00980016_m1
LOC404266	Post GWAS analysis	Hs00420340_m1
HOXB7	Post GWAS analysis	Hs00270131_m1
HOXB8	Post GWAS analysis	Hs00256885_m1
HOXB9	Post GWAS analysis	Hs00256886_m1
PRAC	Post GWAS analysis	Hs00741542_g1
HOXB13	Post GWAS analysis	Hs00197189_m1
TTLL6	Post GWAS analysis	Hs00403596_m1
SIN3B	Post GWAS analysis	Hs00391562_m1
CPAMD8	Post GWAS analysis	Hs00610855_m1
HAUS8	Post GWAS analysis	Hs00928622_m1
MYO9B	Post GWAS analysis	Hs00188109_m1
USE1	Post GWAS analysis	Hs00218426_m1
OCEL1	Post GWAS analysis	Hs00226198_m1

NR2F6	Post GWAS analysis	Hs00172870_m1
USHBP1	Post GWAS analysis	Hs00230579_m1
MERIT40	Post GWAS analysis	Hs00204343_m1
ANKRD41	Post GWAS analysis	Hs00401326_m1
ABHD8	Post GWAS analysis	Hs00225984_m1
DDA1	Post GWAS analysis	Hs00911895_g1
MRPL34	Post GWAS analysis	Hs04194538_sH
TMEM16H	Post GWAS analysis	Hs00394261_m1
GTPBP3	Post GWAS analysis	Hs00378443_m1
CCNE1	MERIT40 functional analysis	Hs01026536_m1
CDKN1A	MERIT40 functional analysis	Hs00355782_m1
CDKN2A	MERIT40 functional analysis	Hs00923894_m1
PTEN	MERIT40 functional analysis	Hs02621230_s1
RB1	MERIT40 functional analysis	Hs01078066_m1
BUBR1	MERIT40 functional analysis	Hs01084828_m1
MAD1	MERIT40 functional analysis	Hs00231137_m1
MAD2L1	MERIT40 functional analysis	Hs01554513_g1
MAD2L2	MERIT40 functional analysis	Hs01057448_m1
CHK1	MERIT40 functional analysis	Hs00967506_m1
CHK2	MERIT40 functional analysis	Hs00200485_m1
BUB1	MERIT40 functional analysis	Hs00177821_m1
BUB3	MERIT40 functional analysis	Hs00190920_m1
ATM	MERIT40 functional analysis	Hs01112307_m1
FANCD2	MERIT40 functional analysis	Hs00276992_m1
PARP1	MERIT40 functional analysis	Hs00242302_m1
RAD51	MERIT40 functional analysis	Hs00153418_m1
RAD51C	MERIT40 functional analysis	Hs00427442_m1
RAD51D	MERIT40 functional analysis	Hs00172529_m1
BARD1	MERIT40 functional analysis	Hs00184427_m1
BRCA1	MERIT40 functional analysis	Hs01556193_m1
BRCA2	MERIT40 functional analysis	Hs00609073_m1
MERIT40	MERIT40 functional analysis	Hs00204343_m1
BRE	MERIT40 functional analysis	Hs01046283_m1
CCDC98	MERIT40 functional analysis	Hs01128826_m1
RAP80	MERIT40 functional analysis	Hs00212459_m1
BRCC3	MERIT40 functional analysis	Hs00827974_m1
ERCC1	MERIT40 functional analysis	Hs01012158_m1
ERCC2	MERIT40 functional analysis	Hs00361161_m1
XPA	MERIT40 functional analysis	Hs00166045_m1
XRCC5	MERIT40 functional analysis	Hs00221707_m1
XRCC6	MERIT40 functional analysis	Hs00995282_g1
POLB	MERIT40 functional analysis	Hs00160263_m1
MLH1	MERIT40 functional analysis	Hs00179866_m1
MSH2	MERIT40 functional analysis	Hs00953523_m1
PMS1	MERIT40 functional analysis	Hs00922262_m1
PIK3CA	MERIT40 functional analysis	Hs00907966_m1
KRAS	MERIT40 functional analysis	Hs00364284_g1
BRAF	MERIT40 functional analysis	Hs00269944_m1
HOXA10	MERIT40 functional analysis	Hs00172012_m1
CDK12	MERIT40 functional analysis	Hs00212914_m1
BAD	MERIT40 functional analysis	Hs00188930_m1
TP53	MERIT40 functional analysis	Hs01034249_m1

ABHD8	MERIT40 functional analysis	Hs00225984_m1
CPAMD8	MERIT40 functional analysis	Hs00610855_m1
DDA1	MERIT40 functional analysis	Hs00911895_g1
GTPBP3	MERIT40 functional analysis	Hs00378443_m1
MRPL34	MERIT40 functional analysis	Hs04194538_sH
MYO9B	MERIT40 functional analysis	Hs00188109_m1
NR2F6	MERIT40 functional analysis	Hs00172870_m1
OCEL1	MERIT40 functional analysis	Hs00226198_m1
USHBP1	MERIT40 functional analysis	Hs00230579_m1
ANKRD41	MERIT40 functional analysis	Hs00401326_m1
TMEM16H	MERIT40 functional analysis	Hs00394261_m1
USE1	MERIT40 functional analysis	Hs00218426_m1
PAX8	FTE project	Hs00247586_m1

Table 2.5: List of all the Taqman® Real time PCR expression assays used for the Fluidigm 96.96 Dynamic array.

2.4.6 Analysis of Real time PCR gene expression data

The expression data obtained from Real time PCR using the ABI7900 Taqman® Sequence detection system or the Fluidigm 96.06 Dynamic array were analysed in a similar manner as follows. Initially, to confirm the linear relationship of expression versus RNA quantity for each target probe and endogenous control probes a standard curve was created using the expression data generated by the serial dilutions of a mix of the cDNAs used. Data were assumed reliable if an excellent fit (Correlation coefficient R^2 between 0.95 and 1) of the standard curve data to a straight line was observed or at least being able to create a suboptimal curve ($R^{2>}0.8$). Additionally, quality control analysis regarding the pass rates for the assays and samples tested was performed to ensure the reliability of the data as will be described in more detail in Chapter 5.

The Real time expression data were analysed using the comparative ΔΔCt method according to Applied Biosystems provided and the expression values of all cell lines were generated relative to either the lowest or highest expression of a NOSE cell line for each cell line normalized against GAPDH and β-actin. To validate the Comparative Ct method of analysis, the data were also analysed using the standard curve method and the relative expressions of each sample were compared between the two analyses to confirm that the relationship would remain the same. Differences in the relative expression of each candidate gene between EOC cell lines and normal cell lines were assessed using the nonparametric two sided Wilcoxon Rank sum test using R

software (R Foundation for Statistical Computing, Vienna, Austria) and P values and representative boxplots using the median of the expression for each group were generated. The following commands were used in R:

To generate P values:

wilcox.test(Relative_gene_expression~Cell_group,alternative="two.sided",paired=FALSE,conf.int=TRUE)

To generate boxplots:

```
Cell_group[Cell_group==2] <- "EOC cell lines (number of samples)"
Cell_group [Cell_group ==1] <- "NOSE & FT cell lines (number of samples)"
those.cell.lines <-as.factor(Cell_group)
levels(those.cell.lines)
pdf(file="NOSE&FTvsEOC_gene.pdf",paper="a4",family="Helvetica")
boxplot(Relative_gene_expression~Cell_group,xlab="Cell lines",
ylab="Relative gene expression",cex.main=0.8, main="Gene expression in NOSE & FT vs EOC cell lines normalised to GAPDH")
```

Statistically significant values were considered if $P<0.05$. In experiments with multiple test multiple testing correction was performed by correcting the P value cut off relative to the statistical tests performed [P= 0.05/(number of statistical tests performed].

2.4.7 Genotype Specific gene expression analysis

The DNAeasy Blood and tissue Kit (Qiagen) was used for DNA extraction of NOSE and FTE cell lines respectively according to the manufacturer's instructions. Extracted DNA was quantified using a NanoDrop® reading the absorbance at 260nm. The ratio of absorbance at 260nm and absorbance at 280nm was used for purity estimation. 250ng of 36 NOSE DNA samples were included in the genotyping performed for the Stage II ovarian cancer GWAS performed by our group and collaborators using a custom Illumina Infinium iSelect BeadChip for a total of 22,790 SNPs including the SNPs of my interest. The genotyping of the 36 NOSE samples was repeated and additional 16 NOSE and 4 FTE cell lines DNA were included in genotyping performed in the context of an association study performed within our laboratory

and collaborators. Briefly, the latter study was (Collaborative Oncological Gene-Environment Study (COGS) where 211,155 SNPs, including the SNPs of my interest, were genotyped in McGill University and Genome Quebec Innovation Centre using a custom Illumina Infinium iSelect assay.

To ensure reliability of the genotyping the minor allele frequency (MAF) of the genotyped SNPs of interest was calculated from the NOSE &FTE samples genotypes and compared to the HapMap-CEU reported MAF. The following formula was used to calculate MAF:

MAF= (2×RH+H/(2×(RH+H+CH)) where
RH: Rare homozygotes, H: Heterozygotes, CH: Common homozygotes.

Genotype specific expression was calculated using expression data previously generated using the comparative ΔΔCt method. The expression of each genotype group was evaluated relative to the average expression of the common homozygotes for each candidate SNP normalised against the expression of the endogenous control genes. To analyse gene specific gene expression, linear regression analysis was initially used to assess the difference in expression between common homozygotes, heterozygotes and rare homozygotes. However the sample size was very small and to avoid the result of regression being skewed from a rare homozygote that is an outlier, the rare homozygotes were combined with the heterozygotes. Wilcoxon Rank sum test was applied to test if the rare allele had a role in the expression of the candidate genes. The statistical analysis and generation of comparative boxplots was performed in R statistical software using the commands previously described.

2.4.8 Genotype specific methylation analysis

One of the studies included in the ovarian cancer GWAS was the UKOPS (United Kingdom Ovarian Population Study) with participants recruited form the UK. There were 691 cases and 1,051 controls included in this study. The UKOPS controls were aged 50-76 years and were healthy postmenopausal females. Germline DNA was extracted at Tepnel, using a chloroform based

extraction method and 800ng (2×400) for each sample. Genotyping data for those samples were available for the SNPs of my interest from the ovarian cancer GWAS. Additionally, methylation data were made available to me by Prof Martin Widschwendter within our group. The methylation analysis was performed on germline DNA of 148 healthy UKOPS individuals using the Illumina Infinium Human Methylation27 BeadChip for approximately 27,000 CpGs mapping to promoters of 14,000 genes in (Teschendorff et al, 2009) and later on additional 108 UKOPS samples (Teschendorff et al, 2010). The methylation status (β-value) of each CpG island associated with the genes of my interest was analysed relative to the genotype of the related SNP. Methylation differences were assessed in two ways. Firstly, methylation status was compared between the common homozygotes and the rare homozygotes/heterozygotes combined using the Wilcoxon rank sum test as previously described to evaluate the effect of the rare allele on methylation. Secondly, linear regression analysis was performed comparing the methylation status of the three genotype groups separately to evaluate the extent of the rare homozygote genotype affecting methylation compared to the other two genotype groups.

Finally, previous reports have shown that methylation status can be affected by age, so the set was divided into three age groups and linear regression analysis was performed in R to determine if methylation of any of the selected CpG islands was affected by age. If no association was detected between the age distribution and the methylation status any positive associations found between methylation and genotype were assumed reliable and non-age biased.

An example of the commands used in R to perform linear regression analysis and representative boxplots were as follows:

To generate P values:

lm(UKOPS_group_ SNP.genotype~Relative_gene_CpG_methylation)

summary(lm(UKOPS_group_ SNP.genotype~Relative_gene_CpG_methylation))

To generate boxplots:

UKOPS _group_ SNP.genotype [UKOPS _group_SNP.genotype ==3] <- "Rare (#)"

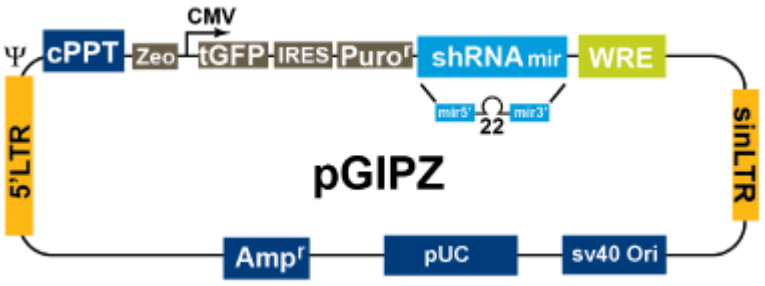
UKOPS _group_ SNP.genotype [UKOPS _group_ SNP.genotype ==2] <-"Rare and hets (#)"

```
UKOPS _group_ SNP.genotype [UKOPS _group_ SNP.genotype ==1] <-
"Common (#)"
those.UKOPS.SNP.genotypes <-as.factor(UKOPS _group_SNP.genotype)
levels(those.UKOPS.genotypes)
pdf(file=" UKOPS_genotype_CpG.pdf",paper="a4",family="Helvetica")
boxplot(Relative_gene_CpG_methylation ~
UKOPS_group_SNP.genotype,xlab=" SNP genotype ", ylab="
Relative_gene_CpG_methylation ",cex.main=0.8, main="Gene CpG
methylation relative to SNP genotype")
```

2.5 *In vitro* assays to investigate functional role of a candidate gene emerging from the GWAS

2.5.1 GIPZ lentiviral shRNA system for gene silencing

The functional role of *MERIT40* was investigated using *in vitro* assays described later on in this section. For this purpose, *MERIT40* was knocked down in EOC cell lines using the lentiviral GIPZ small hairpin RNA (shRNAmir) system purchased from Open Biosystems. This system was chosen to achieve stable knockdown of the gene which is facilitated by the lentiviral delivered shRNA in the infected cell line through the endogenous miRNA pathway. Using lentiviral delivery systems is advantageous compared to retroviruses because they have the ability to infect non-diving cells. The desired shRNAs are cloned into a GIPZ plasmid shown in Figure 2.2.



Plasmid Element	Description
CMV Promoter	RNA Polymerase II promoter
cPPT	Central Polypurine tract helps translocation into the nucleus of non-dividing cells
WRE	Enhances the stability and translation of transcripts
turbo GFP	Marker to track shRNA expression
Puro^r	Mammalian selectable marker
AMP^r	Ampicillin bacterial selectable marker
5'LTR	5' long terminal repeat
pUC ori	High copy replication and maintenance in *e.coli*
SIN-LTR	3' Self inactivating long terminal repeat
RRE	Rev response element
ZEO^r	Bacterial selectable marker

Figure 2.2: Lentiviral GIPZ plasmid containing the shRNAmir sequences

2.5.2 Human GIPZ plasmids containing shRNAmir

Three different *MERIT40* targeting shRNAmir clones were available in the Open Biosystems (OB) library with the distinct sequences shown below.

***MERIT40* shRNAmir 1 (shRNA_M1) (5'-3')** (OB, V2LHS_213618)
TCCGAGGCAGTAGGCA**CTCCTAGTTTGTCATGGATAATTACATCTGTGGCTTCAC TAATTATCCATGACAAACTAGGA**TCGCTCACTGTCAACAGCA

***MERIT40* shRNA mir 2 (shRNA_M2_) (5'-3') (**OB, V2LHS_252352)
TCCGAGGCAGTAGGCA**CCCTCAATGTCTCCCAGAAGATTACATCTGTGGCTTCA CTAATCTTCTGGGAGACATTGAGG** TCGCTCACTGTCAACAGCA

***MERIT40* shRNAmir 3 (shRNA_M3) M3 (5'-3')** (OB, V2LHS_214050)
TCCGAGGCAGTAGGCA**CACCTTCAATCTGGAAGGACTTTACATCTGTGGCTTCA CTAAAGTCCTTCCAGATTGAAGGT**TCGCTCACTGTCAACAGCA

Additionally, three GIPZ plasmids were used as controls. The first one contained a shRNA of non-targeting sequence that would serve as a negative control (OB, RHS4346). Another negative control included was an empty GIPZ plasmid containing no shRNA (OB, RHS4349). As a positive control a GIPZ-shRNA_GAPDH was used (OB, RHS4371).

Finally, for lentivirus packaging, two plasmids were used that contain the genetic information needed for production of the lentiviral particles that will encase and deliver the GIPZ-shRNAs to the EOC cell lines used as well as a gene providing resistance to ampicillin for selection of bacterial colonies. The packaging vectors used were pMDG and p8.91.

2.5.3 Preparation of GIPZ-shRNA and packaging plasmid DNA

All plasmids were provided in DH5α bacterial glycerol stocks. The GIPZ and packaging plasmid containing bacteria were streaked on LB-agar plates containing 100µg/ml ampicillin (Sigma). LB agar plates were prepared using the following recipe: 5g Tryptone 2.5g NaCl, 2.5g Yeast Extract 1.5 % Agar (All Sigma), distilled H_2O to 500ml. The solution was then sterilised by autoclaving and allowed to cool to below 60^{o}C before adding ampicillin (dissolved in 70% ethanol). The LB-agar with antibiotic was then poured into sterile 100mm petri dishes within a laminar flow hood, and allowed to cool and set before replacing

the lid. Once the LB-plates were set the bacterial swabs were spread onto them using the quadrant streaking technique to achieve well-spaced bacterial colonies. Plates were incubated inverted overnight at 37^{o}C. Single colonies were picked from each plate with sterile forceps and inoculated in 3ml of LB medium with 100µg/ml ampicillin depending on the plasmid contained. The medium was then incubated overnight in a shaking incubator with shaking at 225rpm at 37^{o}C. The following day 200µl of the bacterial culture solution was further inoculated in 100ml of LB medium with ampicillin and incubated overnight in a shaking incubator with shaking at 225rpm at 37^{o}C. The following day the bacterial cultures were split in 12ml aliquots and pelleted by centrifugation at 5,200 rpm. The supernatant was then removed and plasmid DNA was prepared using the QIAprep Spin Miniprep Plasmid DNA Isolation kit according to the protocol provided by the manufacturer. The resulting DNA pellets were resuspended in 60µl 10mM Tris-Cl pH 8.5 buffer included in the kit. Plasmid DNA concentration and purity was determined using a NanoDrop® Spectrophotometer.

2.5.4 Sequencing of GIPZ-shRNA plasmid DNA

Following preparation of plasmid DNA all GIPZ plasmids were sequenced to verify the shRNA sequences were correct and contained no mutations acquired during the preparation process (sequencing was performed by Open Biosystems). The sequence for all plasmids was confirmed as expected apart from shRNA_M2 which had a base substitution in which may affect the efficiency of the *MERIT40* knockdown for this shRNA.

2.5.5 Lentivirus GIPZ-shRNA production

The GIPZ-shRNA lentiviruses were produced in 293T packaging cell lines. The procedure was performed as follows. Day one: 293T cells were seeded 24 hours before transfection into a 100mm plate with 8ml complete media (DMEM, 15% FCS, L-Glutamine) and incubated at 37°C, 5% CO_2. 293T cells had to be sub-confluent in the day of transfection. Day 2: The cells were co-transfected with GIPZ-shRNA (or GIPZ_control) plasmid and the lentiviral

packaging vectors pMDG and p8.91. Plasmid DNA was diluted in 1.5ml tubes to a total volume of 15µl with TE pH8.0 as follows: 1µg p8.91 (gag-pol expressor), 1µg pMDG (VSV-G expressor) and1.5µg pGIPz DNA and finally sterile TE added to a final volume of 15µl. For each transfection mix, 200µl of Optimem (Gibco) was added to a second micro-tube and 10µl of Fugene® transfection reagent (Promega) was added to the center of the tube caring not to touch the sides. The mixture was mixed gently by inversion and not vortexed. DNA was then added to the Optimem-Fugene mix and once again mixed by inversion. The mix was centrifuged for 2 seconds to remove the mix from the sides of the tube and incubated at room temperature for 15 minutes. During this incubation, the media on the cells was replaced with fresh pre-warmed complete media. After the 15 minute incubation, the Optimem/DNA/Fugene mix was added to the cells drop wise and the plates were swirled to mix well. The cultures were incubated at 37°C, 5% CO_2 for 24hours. Day 3: The medium containing the transfection mix was replaced with 8ml of fresh complete medium and the cells were returned to 37°C, 5% CO_2. Days 4 to 6: The cells were visualised under a fluorescent microscope to confirm GFP expression The supernatants-containing lentivirus were harvested 24, 48 and 72 hours after the media change that was performed at day 3. The medium of each virus was collected into a 50ml syringe and filtered through a 0.45µm filter (Millipore) into a centrifuge tube. Fresh complete media (8ml) was added to the cells for supernatant to be collected the following day if needed. The viral supernatant was stored at -80^{o}C in 1-3ml aliquots. The transfection procedure and collection of the virus is shown at Figure 2.3.

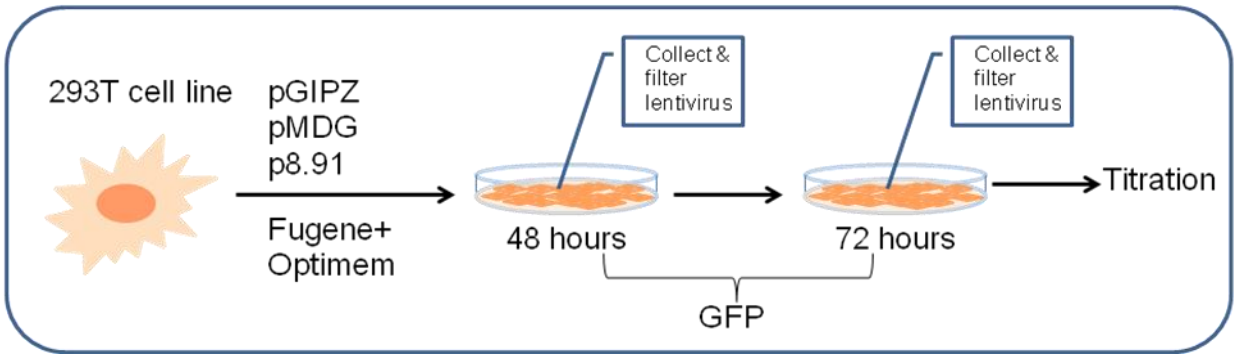


Figure 2.3**:** Lentivirus GIPZ-shRNA production procedure

2.5.6 Lentivirus GIPZ-shRNA titration

To ensure that the lentiviral stock produced was viable and to test what fraction of target cells could be transduced a titration was performed. This enables the number of copies of viral construct per target cell to be controlled. The titration of a virus was performed by placing dilutions of the virus onto a chosen cell line. Several different cells can be used to determine the virus titre but titres may vary to several orders of magnitude thus I decided it was appropriate to use an EOC cell line since the cell lines to be transduced were all EOC cell lines. Fluorescence activated cell sorting (FACS) was used to measure GFP expression of cells and quantifying the amount of lentiviral particles in order to calculate the lentivirus transducing units (TU) per ml.

The day before titration, 100,000 C13 EOC cells were seeded into all wells of a 6-well tissue culture plate, ensuring a uniform spread of cells on the bottom of the wells. One 6-well plate was prepared for each lentiviral stock to be titrated. One control plate of uninfected cells was plated as a negative. The cells were incubated at 37°C, 5% CO_2 for a couple of hours until attached. 1:3 serial dilutions of the viral supernatant containing 8µg/ml polybrene were prepared and added to 6 wells as shown at Figure 2.4. Polybrene (Sigma) is a cationic polymer which is used to enhance the efficiency of lentiviral infection of cells. The infected cells were incubated for 48-72 hours at 37°C, 5% CO_2.

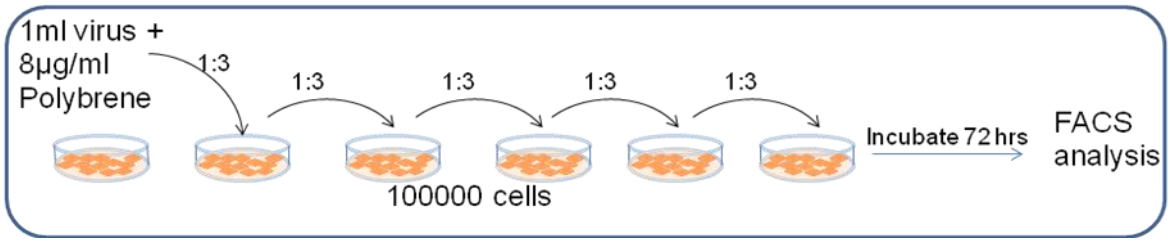


Figure 2.4: Lentivirus GIPZ-shRNA titration procedure.

Before fluorescence-activated cell sorting (FACS) analysis, the culture medium of the infected and uninfected control was removed and the cells washed with PBS and trypsinised. The cells were resuspended thoroughly to disrupt clumps and then transferred to a FACS tube (ELKAY) containing 500µl PBS. If the FACS analysis was not performed within 1 hour, cells were fixed in 4% (w/v) paraformaldehyde solution for 30 minutes and kept for maximum one

week at 4°C. The FACS analysis was performed in a CyAn ADP flow cytometer (Dako Cytomation).The percentage of GFP-positive cells was then determined by FACS analysis. The titre was then calculated using the dilutions that had 10-20% GFP positive cells. The titre of the virus is calculated in transducing units (TU)/ml, according to the formula:

$$\frac{(1\times10^5 \text{ seeded cells}\times \%\text{ GFP-positive cells})}{\mu\text{l of vector}}$$

2.5.7 Stable infection of EOC cell lines with lentivirus GIPZ-shRNA

The EOC cell cultures to be infected were trypsinised and counted as previously described. 1×10^5 cells were plated in their corresponding complete media in a 100mm plate and incubated at 37°C, 5% CO_2 for 2-4 hours until they attached. The media was then removed and replaced with 10ml of complete media containing 8µg/ml polybrene and the appropriate amount of virus supernatant. The amount (µl) of virus supernatant to be added was calculated based on the TU established for each virus at the titration using the following formula:

$$\frac{(1\times10^5 \text{ seeded cells}\times 1\text{ MOI})}{\text{TU/ml}}$$

MOI stands for the multiplicity of infection which determines the amount of transducing units per cell. 1TU/cell was used. The TU/ml calculated for each virus and the amount of supernatant used based on the desired MOI of 1 was as follows:

Lentivirus GIPZ-shRNA	TU/ml	MOI	Cells to infect	Virus supernatant to infect with (µl)
GAPDH	3.56×10^5	1	1×10^5	281
Empty	4.40×10^5	1	1×10^5	228
Non-Silencing	3.93×10^5	1	1×10^5	255
MERIT40 (M1)	9.37×10^5	1	1×10^5	1067
MERIT40 (M2)	1.89×10^5	1	1×10^5	5284
MERIT40 (M3)	1.56×10^5	1	1×10^5	639

The infected cells were incubated for 72 hours and subjected to puromycin selection. Infected cells were selected with 0.5-1.5 µg/ml puromycin,

depending on the cell line, to avoid any contamination of the stably infected cell lines with non-infected cells. Alongside, control uninfected cell lines were seeded and treated with puromycin. The control cell lines were 100% killed in 3-7 days.

2.5.8 Confirmation of *MERIT40* knockdown

2.5.8.1 Confirmation of *MERIT40* knockdown at the mRNA level

RNA was extracted from the stably infected and the relative control EOC cell lines and transcribed into cDNA as previously described. Gene expression was determined by Taqman® Real time PCR probes in ABI7900 Taqman® Sequence detection system and analysed with the comparative ΔΔCt method as previously described. The samples used as calibrators were the control EOC cell line for each set of infected cell lines.

2.5.8.2 Confirmation of *MERIT40* knockdown at the protein level

Cell line protein extraction: The transduced cell lines were cultured in 145mm plates until they reached 70% confluence. The cells were then starved for 24 hours using serum free culture media in order to reach basal levels of protein. The plates were then placed on ice and washed twice with ice cold PBS. 500µl of lysis buffer (0.02% NP40, 6mM Hepes pH 7.4, 30mM NaCl, 0.2mM EDTA, 17µg/ml aprotinin,1µg/ml pepstatin, 1µg/ml leupeptin, 100µg/ml AEBSF and phosphatase inhibitors cocktail) (all Sigma) was added on plate and the cells were scraped with a cell scraper while on ice. The cell lysate was then quantified or stored at -20^{o}C until used. Before quantification cell debris were cleared by centrifugation at 13,000 rpm for 20min at 4^{o}C. The supernatant was then transferred to a fresh tube and the pelleted cell debris discarded. The protein lysates were then quantified diluted to 1.25µg/ml and 5X Laemli buffer (50mM Tris pH 6.8, 10% v/v glycerol, 2% SDS (w/v), 0.1% bromophenol blue (w/v), 2% beta-mercaptoethanol) (all Sigma) was added to give a final protein

concentration of 1µg/ml. Before use, protein samples were denatured at 100^{o}C for 3min.

Bradford assay for protein quantification: Protein samples were quantified using Bradford assay and the Commassie protein assay kit (Pierce). The Bovine Serum albumin (BSA) standards were prepared as described in the manufacturer's protocol and the proteins were quantified according to the standard microplate protocol provided by the manufacturer. 2µl of each standard or unknown was used instead of the 5µl mentioned in manufacturer's instructions. 1:2 dilutions of the unknown concentration proteins were used. After spectrophotometric measurement at 595nm, a standard curve was generated and the concentrations of the unknown protein samples were extrapolated from the standard curve's equation, accounting for the dilution factor as appropriate.

Western Blot: The protein lysates of the control and infected cell lines were shipped to University of Southern California, Keck School of Medicine, Department of Preventive Medicine where Dr Kate Lawrenson performed a western blot with a custom made antibody that was provided by Chen J.'s laboratory, Department of Therapeutic Radiology, Yale University School of Medicine (Feng et al, 2009). Ponceau staining was used as loading control.

2.5.9 Measuring cell viability using MTT

The viability of cell lines following inducible stress and also their ability to proliferate in culture was evaluated by performing an MTT assay (3-(4,5-Dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide, a yellow tetrazole) colorimetric assay. MTT is reduced by the mitochondria of living cells to a purple formazan product.

MTT assay optimization: The cell number that should be plated for assays that would be followed by MTT end point readouts and MTT reduction is cell type dependent and was optimised to avoid surpassing the limit of detection. The optimal MTT incubation time was also determined for the different cell lines. I have tested the intensity of the absorbance reading after plating 2000-12000 cells and have created graphs for the number of cells

versus the absorbance's intensity for the EOC cell lines tested. All the cell lines showed a linear relationship of cell number versus intensity. According to the MTT optimization performed 6000-10000 cells were plated in the subsequent MTT assays for cell lines SCOV3IP, EFO27, HOC7, TOV112D, MPSC1, OAW42. The MTT incubation times were found to be optimal between 2.5 and 3.5 hours. Plating 6000-10000 cells and incubated MTT between 2.5-3.5 hours was also acceptable for OVCAR3, A2780/A2780CP, OV2008/C13 cell lines, according to previous optimisation experiments performed for these cell lines within our group (Miss Sheetal Dyal, personal communication).

MTT assay for viability or proliferation: The appropriate amount of cells was seeded in 96 well tissue culture plates and incubated at 37°C, 5% CO_2 for 72 hours. The assays were performed in 96 well plates and all samples were run in 8 replicates. A 1mg/ml MTT (Sigma) solution was used. The culture medium was aspirated from the cells and 100µl of MTT was added per well. After 2.5-3.5 hours incubation 37°C, 5% CO_2 the MTT was removed and 100µl of dimethyl sulfoxide (DMSO) (Sigma) was added to solubilize the MTT reduced by the mitochondria of the cells producing a purple colour. The absorbance that was indicative of cell viability/proliferation was then read at 562nm using the Tecan Spectrafluor Plus plate reader.

2.5.10 Cisplatin and carboplatin dose response assays

6000 to 10000 cells were plated with 100µl of the appropriate medium in each well of a 96 well plate 16 hours before dosing with cisplatin (UCLH pharmacy) or carboplatin (UCLH pharmacy) leaving 4 wells empty for background control. The seeded cells were incubated 37°C, 5% CO_2. After 16 hours incubation the culture medium was removed and replaced with fresh complete medium. The assay was performed in quadruplicate for each cisplatin concentration. The cisplatin and carboplatin concentrations that were used for the dosing are shown at Table 2.6. The cisplatin and carboplatin dilutions were then prepared to yield final concentrations in 200µl. The cells were then incubated for 48hours. The medium containing cisplatin was then removed and an MTT assay was performed as described previously to assess cell viability.

The dose response curves and IC50 (amount of drug needed to kill 50% of the cells) of the cell lines to cisplatin and carboplatin were generated by PRISM software, after performing a nonlinear regression analysis for the production of a variable slope of cisplatin concentration versus cell viability and the 95% Confidence Intervals (CI) were also calculated.

2.5.11 Gamma H2AX immunofluorescence staining

Gamma H2AX staining was performed in cells cultured in normal conditions in order to assay the ability to repair spontaneous endogenous DNA damage**.** Round coverslips were placed on a 6 well plate. Cells to be assayed were trypsinised, pelleted and counted. 2×10^5 cells were seeded per well and incubated at 37^{o}C, 5% CO_2 for 48 hours. After incubation cells were fixed on the coverslips with 4% paraformaldehyde (Sigma). Fixed cells were incubated overnight at 4^{o}C in PBS. Cells were stained as previously described in the immunofluorescence cytochemistry section using a 1:1000 dilution of the primary γH2AX antibody (Abcam). Pictures of the stained cells were obtained under the 40x objective of an Olympus BX64 fluorescence microscope. 50 cells per condition were monitored for expression and intensity of γH2AX which indicated active DNA repair following spontaneous DNA damage and the results were plotted in histograms.

2.5.12 Irradiation of cells to assay DNA repair for exogenous DNA damage

X-Ray irradiation of cells was performed to account for the ability of cells to repair the DNA damage induced. 8×10^4 cells were plated in complete media in a 96-well tissue culture plate leaving 4 empty wells to be used as background control. The cells were incubated at 37^{o}C, 5% CO_2 overnight. The following day, the cells were subjected to increasing doses (1, 10 and 100 Grays) of ionising radiation (X-Ray). Following irradiation the cells were incubated for 72 hours at 37^{o}C, 5% CO_2. The ability of cells to repair damage and survive was assayed by performing an MTT assay and generating dose response curves and the IC50 of the cells to irradiation dosing as previously described.

Carboplatin Concentrations (μM)	Cisplatin Concentrations (μM)
0	0
0.3	0.3
0.5	0.5
1	1
3	3
5	5
7.5	10
10	15
12	20
14	30
16	40
18	50
20	60
25	70
30	80
40	90
50	100
70	120
90	140
110	160
130	180
150	200

Table 2.6: Cisplatin and Carboplatin concentrations used for dosing EOC cells to generate dose response curves.

2.5.13 Cell cycle analysis using flow cytometry

Cultured cells of no more than 80% confluency were trypsinised, centrifuged, washed in PBS and resuspended in 1ml of 70% ethanol in serum free media. The cells were left overnight at 4^{o}C to be fixed. The cells were then centrifuged at 4,000rpm in a microcentrifuge for 2 minutes. The cell pellet was resuspended in 0.5-1ml PBS containing 50µg/ml propidium iodide (PI) (Sigma) which stains the cell's DNA and 40µg/ml RNase A to get rid of any RNA traces that can bind PI. The samples were left for 30 minutes in the dark at room temperature. The cell cycle profiles were generated using a CyAn ADP flow cytometer (Dako Cytomation) using the FL3 channel (>580nm). Cell cycle profile analysis was performed in Summit.v4.3 software (Beckman Coulter)

2.5.14 Proliferation analysis with BrDU labelling using flow cytometry

Cultured cells were allowed to reach 80% confluency. The cells were removed from the incubator and BrDU was added to each plate to a final concentration of 10µM. One plate for each cell line to be tested was left with no added BrDU to serve as a negative control. The plates were incubated for 1 hour at 37^{o}C, 5% CO_2. The media was then aspirated and the cells trypsinised, washed twice with PBS and fixed with 0.5-1ml cold ethanol at 4^{o}C overnight. The following day the cell pellet was washed twice with PBS and then resuspended in 2M HCL/5%Triton X-100 (Sigma) adding dropwise while vortexing. The cells were then incubated for 30 minutes at room temperature. The acid was spun off and the pellet washed twice PBS and once with PBS-T (PBS, 0.1% BSA, 0.2% Tween 20, pH 7.4) (all Sigma). 2µl of primary anti-BrDU antibody (Becton Dikinson) was added directly to the cell pellet and incubated for 20minutes at room temperature in the dark. Following incubation the pellet was washed twice with PBS-T and incubated with 50µl FITC-conjugated rabbit anti-mouse (Fab') 2 fragments (DAKO) secondary antibody in a 1:10 dilution for 20 minutes at room temperature in the dark. Finally, the cells were washed with PBS, resuspended in 300-600µl of PBS containing 50µg/ml propidium iodide (PI) (Sigma) and 40µg/ml RNase A and incubated at room temperature in the dark for 20 minutes. The analysis was performed by flow cytometry using a CyAn ADP flow cytometer using the FL3 channel for PI (>580nm) and the FL1 channel for FITC (515-545nm). FACS profile analysis was performed in Summit.v4.3 software.

2.5.15 Preparation of metaphase spreads

Slide cleaning: Microscope slides (Chance Propper, gold star washed) were washed in slide racks in a solution of 3% conc. HCl in 70% methanol (Analar): and 30% distilled water for a minimum of 1 hour but not more than 24 hours. They were then rinsed with 10 changes of distilled water (2 litres per wash).

Preparation of solutions: The following solutions were prepared: Hypotonic solution 0.075M KCl (pre warmed at 37 C) and Carnoy's Fixative consisting of 3:1 Methanol and Acetic acid (Sigma). Carnoy's fixative was used only freshly prepared.

Arresting cells in metaphase and preparing metaphase spreads: Cells were cultured in 100mm plates to 85% confluence. Colcemid (Sigma) was added to cell cultures to a final concentration of 0.1µg/ml. The cell cultures were then incubated for 2 hours and 15 minutes at 37°C, 5% CO_2. Culture media containing colcemid was removed and cells washed with PBS, trypsinised and centrifuged for 5 minutes at 1500rpm. The supernatant was removed and pellet flicked loose.1.5 ml of pre-warmed hypotonic solution was then added slowly against the side of the tube while gently vortexing and the volume was then brought up to 6ml. The cells were incubated at 37°C for 20min. Following incubation the tube was inverted to mix. The reaction was stopped by adding 10 drops of fixative while gently vortexing. Cells were spun down and supernatant discarded. The pellet was flicked and 1.5ml of freshly prepared fixative was added slowly against the side of the tube while gently vortexing. The volume was brought to 6ml and tube was inverted to mix before incubation at room temperature for 10min. The tube was again inverted to mix, cells pelleted and the supernatant removed. The pellet was flicked and 1.5ml of freshly prepared fixative was slowly added against the side of the tube while gently vortexing and the volume brought to 6ml, tube inverted to mix and incubated at room temperature for 30 minutes. Cells were centrifuged, supernatant discarded and 2ml of freshly prepared fixative was added dropwise while vortexing..

Cell dropping procedure and metaphase spreads visualisation under the fluorescent microscope: Cells were dropped using a Pasteur pipette onto the pre-cleaned slides from approximately 3-5 cm height. The slides were allowed to dry before 10ul of 10mg/ml DAPI was added on the slide and a coverslip placed and positioned on top. The metaphase spreads were then visualised under an Olympus BX64 fluorescence microscope under the 100× oil objective and chromosomes of 4-10 cells per sample were counted.

2.5.16 Migration assays

Each condition was run in triplicate. Cells were cultured in a 100mm plate to not more than 80% confluency. Cells were then trypsinised, pelleted and washed in PBS to remove traces of serum. Cells were then resuspended in 20ml serum free media (SFM) and counted. A cell suspension was then prepared in SFM to a final concentration of $6x10^4$ cells/ml. 750µl of culture media supplemented with 10% FBS, used as chemoattractant, was added in the wells of a 24 well plate. 24-well inserts (ThinCert 8µm, Greiner) were transferred with sterile forceps on top of the wells containing the chemoattractant taking care not to trap bubble beneath the membranes. Immediately after placing the insert on the chemoattractant, 0.5ml of the cell suspension prepared ($3x10^4$ cells per well) was added to the inserts. The plates were then incubated for 20 hours at 37^oC, 5% CO_2. Non migrating cells were removed from the upper surface of the membrane by insertion of a cotton swab into the insert and applying gentle pressure while moving the tip over the membrane surface. To stain the migrated cells 1.5ml of 1% Crystal Violet (Sigma) was added in the wells of a clean 24 well plate. The inserts were dipped in the stain for 15min then washed in the PBS for 15 min. Excess liquid and stain was removed from the inside of the inserts using a cotton swab. The inserts were placed upside down and the membranes allowed to dry overnight. The membranes were then carefully removed from the inserts using a scalpel and placed bottom side down on microscope slides containing a drop of immersion oil. A second drop of oil was then placed on top of the membrane and a coverslip was placed on top of the membrane and gentle pressure applied to remove any air bubbles.

The migrating cells were observed and photographed under 40x magnification of a light microscope. The number of migrated cells were counted using cell counter plug-in in Image J software (Schneider et al, 2012). Ideally cells in 4 central fields of triplicate membranes were counted. Histograms of mean number of migrating cells were plotted for the controls and treated cell lines with error bars indicating the variation between replicates.

2.5.17 Assay for Anchorage Independent Growth

This assay was performed in 6-well plates. 6.0% agar containing bacto-peptone (all Sigma) at 10mg/ml was melted at 65°C and added to pre-warmed complete medium in a 1:10 ratio. The solution was immediately thoroughly mixed and 1.6 ml was quickly dispensed into each well of a 12 well tissue culture plate. Agar was allowed to set and used on the same day. The cells to be tested for anchorage independent growth were trypsinised and resuspended thoroughly in complete media to obtain a single-cell suspension and counted. Complete medium was then warmed to 37^{o}C and appropriate number of cells added to the medium. Each cell line was plated in triplicate in numbers ranging from 500 to 10000 cells depending on the experiment. A positive control EOC cell line, TOV112D, was used to test that the gel heat did not kill the cells within the suspension when plated. One tenth volume of 3.3% melted agar (containing 10mg/ml bacto-peptone) was added immediately to the cell suspension, mixed well and 1.2ml was gently overlayed on the base layer. Agar was allowed to set at room temperature and incubated at 37^{o}C, 5% CO_2 for 2-4 weeks depending on the cell line's doubling times. After 2-4 weeks, the agars were fixed and colonies stained in red/purple with 0.5ml p-iodotertazolium violet (1mg/ml), prepared in absolute methanol 24 hours at 37^{o}C. Pictures of the soft agar plates were obtained on a Box gel documentation system using the GeneSnap image acquisition software (Syngene) and the colonies were counted using the cell counter plug-in in Image J software.

3 Establishing *in vitro* models of primary normal ovarian surface epithelial and fallopian tube epithelial cell lines.

3.1 Introduction

EOC is a disease characterised by heterogeneity and one of the major challenges for studying and elucidating the molecular events that leads to the development of the distinct EOC subtypes has been the identification of the site of origin for each of those subtypes. Although the ovarian surface epithelium (OSE) has been conventionally thought to be the site of origin for all EOCs, there is emerging evidence that a big proportion of high grade serous carcinomas emerge from FTE cells (Piek et al, 2003, Medeiros et al, 2006, Tone et al, 2008, Callahan et al, 2007). More recent evidence support that the secretory cells of the fallopian tube (FTSEC) are the sole cell of origin for both low grade and high grade serous carcinomas (Kurman et al, 2011). FTE cells have also been proposed as potential precursor cells of poorly differentiated mucinous adenocarcinoma (Shan et al, 2012, *in print*).

There are growing data to support an FTE & FTSEC origin for a proportion of high-grade serous EOCs, however there are no studies to date that have been able to provide evidence to exclude the OSE as an origin for this subtype. In contrast, there is plenty molecular evidence that supports the OSE origin of several EOC subtypes as previously discussed. This evidence include reports of p53 signatures in OSE lined inclusion cysts of *BRCA1* and *BRCA2* mutation carriers and reports showing that expression of distinct *HOX* genes in OSE are giving rise to distinct EOC subtypes (Naora, 2007, Pothuri et al, 2010). Our group proposes a model supporting a multiplicity of sites of origin for EOC and suggesting that different subtypes emerge mainly either from the normal ovarian surface epithelium or from the distal end of the fallopian tube (Kindelberger et al, 2007). Thus, it is essential to study the pathogenesis of EOC using both NOSE and FTE *in vitro* models.

Normal ovarian surface epithelial (NOSE) cells isolation (Auersperg et al, 1984), culturing conditions (Li et al, 2004) and immortalisation (Maeda et al, 2005) have been well established and optimised over the past decades. NOSE cells have successfully been immortalised and used in models to study transformation monitoring their phenotype after induction of genes such as *TNFα*, *MYC*, *BRAF*, *KRAS* (Maeda et al, 2005, Kwong et al, 2009, Lawrenson et al, 2011)

The research in order to isolate FTE cells has increased over the last few years since their proposed role of being the precursors of certain types of EOC. Earlier studies have tried to isolate epithelial cells of the fallopian mucosa using methods such as enzymatic treatment with trypsin and/or pancreatin. The epithelial status of the isolated cells was confirmed with anti-cytokeratin antibodies of CK7, CK8, CK18, CK19 and PKK1. The lifespan of the isolated cells was reported to be extremely short and all studies have failed to successfully subculture FT isolated cells (Henriksen et al, 1990, Comer et al, 1998). More recent advances have isolated FTSEC cells and cultured them *ex vivo* on collagen matrix (Levanon et al, 2010) but did not sub-culture them or compared their phenotypic characteristics with FT tissue *in vivo*.

The classical techniques for isolating and culturing epithelial cells involve culturing of cell monolayers on plastic surfaces. These traditional two-dimensional (2D) *in vitro* techniques are causing the loss of the normal architecture, geometrical features and cell-cell interactions and interactions of cells with extracellular matrix that characterise tissues *in vivo*. Seminal work in three-dimensional (3D) modelling by Bissell and colleagues has shown that 3D culturing of normal breast epithelial cells can induce gland formation, restore cellular polarity and induce up-regulated expression of biologically active molecules expressed *in vivo* (Aggeler et al, 1991, Jones et al, 1993, Streuli et al, 1995). Recent work within our group has generated 3D models of NOSE cells in order to better model the microenvironment of the ovary *in vivo*. The morphological and biological characteristics have been compared between NOSE 3D and 2D models and OSE *in vivo* to confirm that 3D cultures more closely resemble the tissue *in vivo* than the 2D cultures. In this study two different techniques for establishing 3D cultures were tested, the polyHEMA

coated vessels and the rotary cell culture system (RCCS). The 3D structures grown in the RCCs had a chaotic internal structure and exhibited higher levels of apoptosis compared to the polyHEMA grown 3D cultures. Thus, use of the polyHEMA coated vessels to culture 3D cultures was proposed to be a more appropriate 2D modelling approach (Lawrenson et al, 2009).

Aims of this chapter:

1. The first large scale NOSE cell line repository (n>50) will be collected and characterised to be further used to study the expression of candidate genes involved in EOC development in following chapters.
2. Collect, establish and sub-culture primary FTE cell lines (n>3).
3. Establish 3D cultures of FTE cell lines. Investigate the morphological and growth and lineage characteristics of the primary FTE cell cultures in 2D and further compare with 3D FTE cultures and FT tissue.
4. The differential expression of genes that are known to have an implication in ovarian cancer development will be assayed between of the established NOSE & FTE cell lines and EOC cell lines using real time expression assays.

3.2 Establishing and characterising a primary NOSE cell line repository

NOSE cell lines were isolated from brushings of normal ovaries during surgical procedures that took place in University College London Hospital (UCLH). The women were identified as suitable for participation in this study when their diagnosis did not involve malignancy or disease of the ovaries and they were not carriers of a familial mutation in any of the high risk genes identified for ovarian cancer subjected to prophylactic salpingo-oophorectomies.

The patients were consented the day before or on the day of surgery. When the surgical procedure was open body surgery, the brushings were collected in surgery after the specimen was removed from the patient by brushing the surface of both the left and the right ovary when possible as described in more detail in the methods section of this project. In the case of laparoscopic surgery, brushings were either collected after excision of the specimen or by the surgeon while the ovaries were still in the patient.

All the specimens that were suitable for this study would have to have no malignancy and this was confirmed in three stages. Firstly, ovaries were confirmed to be normal on the day of the surgery based on the morphology of the ovaries by the surgeon in charge. Secondly, a report was drawn by a pathologist (Dr Rupali, Dr Benjamin, The Rockfeller Building, UCLH) after histopathological examination of cross sections of the samples. Finally, malignancy was tested by fluorescent immunocytochemical staining of the established cell lines for the ovarian cancer marker CA125.

The NOSE brushings collected were immediately transferred to NOSE culturing media (NOSE-CM). This is a culture medium that contains insulin, epidermal growth factor (EGF), bovine pituitary extract (BPE), hydrocortisone (HC) and foetal bovine serum (FBS) (Li et al, 2004).

Growth of epithelial colonies was observed under the microscope within 5-12 days following initial seeding. The morphology of the colonies in the initial seeding flask was cuboidal, and as soon as they were further passaged became more elongated, fibroblast like resembling normal epithelial-mesenchymal transition (EMT). Studies have previously shown that epithelial cells can undergo EMT *in vitro* coupled with over-expression of FSP and over-

expression of several cytokines can contribute to this, one of the most important being the EGF (Okada et al, 1997). The optimal culturing medium for extending the *in vitro* lifespan of NOSE cell lines contains EGF which likely contributes to the observed EMT.

A total number of 72 patients were consented and subjected to isolation of brushings, with brushings from the left and right ovary collected from 53 patients, brushings solely from the right from 10 and brushings solely from the left from 8 patients. A total of 124 brushings were collected from those 71 patients. The success rate for NOSE collection with this protocol is as high as 88% as 6 brushings (from 3 patients) were lost due to bacterial contamination, and 9 (from 5 patients) did not grow. Of the 63 remaining patients, 3 of them were reported to have ovarian malignancies after histopathological examination of the cross sections and were also excluded from this study. 60 NOSE cell lines were isolated and established as suitable to use in this study (six of them were contributed by Dr Kate Lawrenson that isolated and established them as part of her PhD). Patient information as well as the histopathological report for each is presented on Appendix 1, Table 1. Previous work in our group in four of the NOSE cell lines has established that their life span ranges between 12- 15 passages. Therefore all the work with the NOSE cell lines has been done whilst they would be up to their 8^{th} passage to ensure that their population doubling times would remain in a linear scale.

The 60 suitable for this study NOSE primary cell lines were cultured and frozen for subsequent passages. The established primary NOSE cell lines were additionally genotyped for 7 SNPs, in order to allow for their positive genetic identification in the future. A table summarising the NOSE genotyping information is shown in Appendix 1, Table 4.

The NOSE cell lines were further characterised by immunofluorescence cytochemistry for the following markers:

AE1:AE3 **:**Pancytokeratin. Keratins are proteins insoluble to water that form the filaments composing part of the cellular cytoskeleton in epithelial tissues. Nineteen human epithelial keratins have been found and can be acidic (Type I) or basic (Type II) (Moll et al., 1982). Acidic and basic keratins are found

together in pairs the composition of those pairs depending in the type of the epithelial cell, its growth environment and stage of differentiation. AE1:AE3 is a pooled pan-cytokeratin antibody with the AE1 part recognising parts of the acidic subfamily and the AE3 part recognising members of the basic subfamily and is a very good general stain for cells with epithelial origin. This marker was used in order to confirm the epithelial status of the NOSE cell lines.

CK7 **:**Cytokeratin 7**.** Ck7 is a basic (type II) keratin and has been found to be expressed in various tissues including the epithelial duct of the genitourinary tract. This marker was also used in order to assay for the epithelial status of the NOSE cell lines.

CA125 **:**Cancer antigen 125. It is a membrane associated mucin found in epithelial cells of the female reproductive track. The main biological function of this protein is to provide a barrier which prevents the attachment of infectious agents and foreign particles on the epithelial cell surface. CA125 levels have been found to be elevated in the patients with invasive ovarian cancer (Osman et al, 2008). This marker was used in order to confirm that the collected cell lines were not malignant.

FSP **:**Fibroblast specific protein. It is a member of the calmodulin S100 troponin C superfamily. FSP is a protein expressed in cells of mesenchymal origin such as fibroblasts and is found in the cytoplasm of fibroblasts but not cells of epithelial origin. However several studies have shown that epithelial cells in early development or even later in adults can undergo a morphological conversion from cuboidal to more elongated fibroblast –like appearance and decreased expression of cytokeratin, a process that is known as EMT. This process has been linked to the over-expression of FSP (Okada et al, 1997). NOSE cell lines have been shown to undergo EMT *in vitro* (Lawrenson et al, 2009). This marker was used in order to confirm that the NOSE cell lines did not have fibroblast contamination from the stromal cells in the ovary. However due to expected EMT low levels of FSP expression present in the established cell lines is acceptable.

FVIII **:**Factor VIII. It is a blood clotting factor and is a glycoprotein which is expressed by endothelial cells. This marker was used to confirm that there was no blood cell contamination within the NOSE cell lines due to the nature of the surgical procedures prior collection.

Each immunofluorescence staining experiment was performed with positive and negative controls. Secondary antibody negative controls were NOSE cell lines that were incubated with only the FITC (Fluorescein isothiocyanate) conjugated secondary antibody without prior incubation with a primary antibody. Positive and negative controls for marker expression were cell lines that from literature are known to express or not a specific marker. As positive control for FVII Human Umbilical Vein Endothelial Cells (HUVEC) were used. None of the NOSE cell lines exhibited any endothelial cell contamination as none of them stained positive for FVIII. T71, an endometrial cancer cell line, was used as a positive control for markers AE1:AE3, CK7 and CA125.

All 60 primary NOSE cell lines that were established stained positive for epithelial markers. The positive expression of epithelial markers was scored as low, moderate and strong depending on the % of cells staining positive (Table 3.1). Based on the large number of the primary cultures established and characterised I propose that the expression of marker AE1: AE3 is a stronger indicator for NOSE epithelial status than CK7. 51 of the characterised cell lines demonstrated strong or moderate expression of AE1:AE3 whereas only 34 cell lines demonstrated strong or moderate expression of CK7. None of the cell lines stained for the cancer antigen CA125 consistent with the reports of the histopathologist for the selected cell lines reporting them as non-malignant. Finally, INOF2 which is a normal ovarian fibroblastic cell line was used as positive control for the FSP marker (Lawrenson et al, 2010). INOF2 had a very strong expression of FSP (Figure 3.2), an indication that the low expression of FSP observed in several NOSE cell lines was due to EMT and not due to stromal contamination. Only one cell line, NOSE 217 L, exhibited a very strong expression of FSP and this cell line was excluded from any further studies. An example of the expression of these markers in three NOSE cell lines and the control cell lines is shown in Figure 3.2 and a summary for the staining of all 61

NOSE cell lines is shown in Table 3.1. The images for the characterisation of all 60 NOSE cell lines are presented in Appendix 1, Figure 1.

An additional observation made was that the expression levels of epithelial markers differed between the patients. In an attempt to investigate how and if pan-cytokeratin expression would modify depending on the age of the patient the mean age of the patients that their OSE brushings stained strong, moderate or low for AE1:AE3 was calculated and is presented in Figure 3.1. I found that there was a slight trend indicating that expression of cytokeratins may decrease as the age of patients increase. However after linear regression analysis in R the result was not statistically significant as the deviations of each group were significantly overlapping.

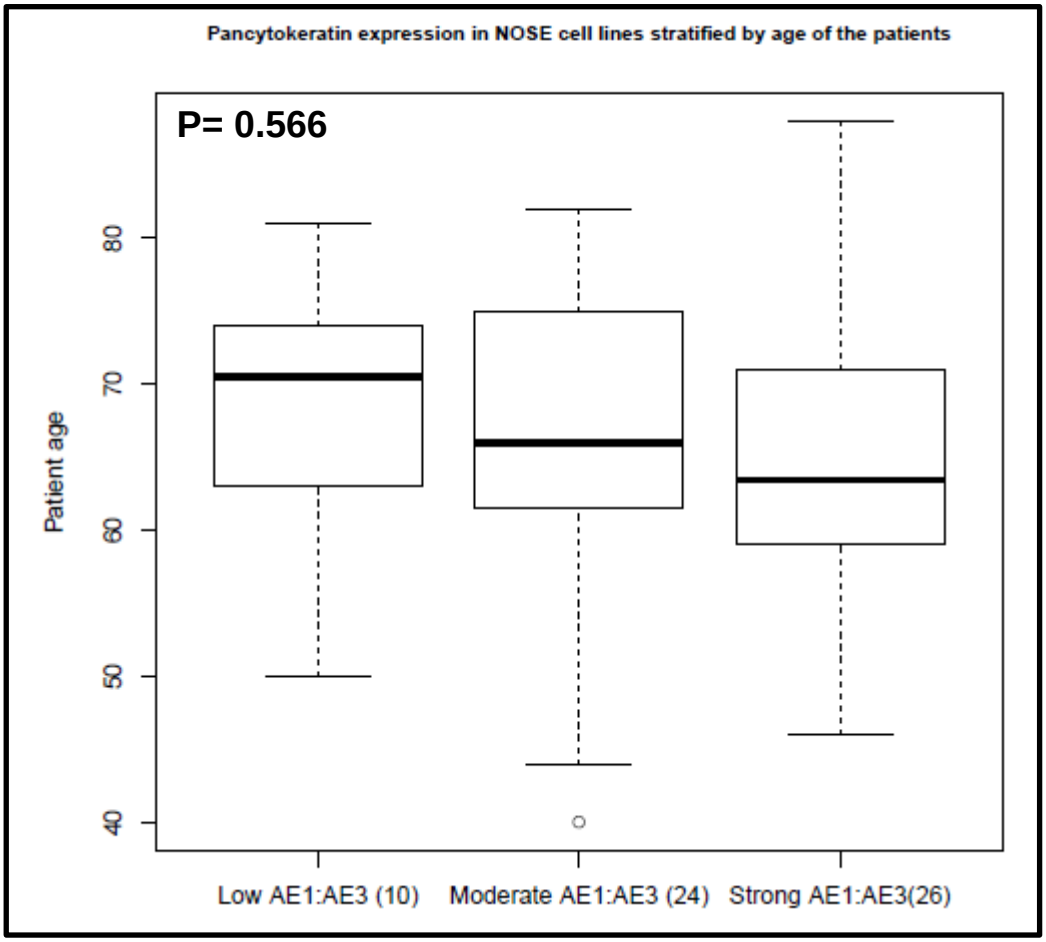


Figure 3.1**:** Correlation of patient age with expression of AE1:AE3 marker for pancytokeratin for isolated NOSE cells.

NOSE cell line	Passage	AE1:AE3	CK7	CA125	FSP	FVIII
3	7	++++	++	-	+/-	-
9	8	++	+	-	+/-	-
11	5	+++	++	-	-	-
13	5	++	+	-	-	-
18 L	7	++	+++	-	-	-
24 R	5	++++	+++	-	-	-
175 L	4	+	-	-	-	-
176 R	6	+++	++	-	+/-	-
179 R	5	+++	+	-	-	-
181 L	5	++	+	-	-	-
185 R	4	+	++	-	+	-
189 L	6	+++	++	-	-	-
191 L	5	+	+	-	-	-
192 L	4	++	++	-	-	-
197 L	4	++++	+++	-	+/-	-
198 L	5	++	+++	-	+/-	-
200 L	6	+++	+	-	+	-
206 L	6	+++	+++	-	+/-	-
211 L	4	+++	+	-	+/-	-
216 L	4	+++	-	-	-	-
217 L	8	+++	+/-	-	++++	-
218 L	6	+++	+++	-	-	-
224 L	4	++++	+++	-	-	-
228 R	3	+++	++	-	-	-
229 L	5	++	++	-	-	-
230 R	5	+++	+	-	+/-	-
231 L	4	+	++++	-	-	-
232 L	5	+++	+++	-	+/-	-
236 L	7	+++	++	-	+/-	-
238 L	5	++	++	-	-	-
239 L	6	++	+	-	-	-
241 L	6	++	+/-	-	+/-	-
243 R	3	+++	+++	-	+/-	-
245 L	4	+++	+++	-	+	-
246 L	5	++	++	-	+/-	-
250 L	4	++	++	-	+/-	-
252 L	4	++	+	-	-	-
253 L	4	+++	++	-	+	-
254 ?	7	+++	++	-	-	-
255 L	4	++	+	-	-	-
257 L	4	++	+	-	-	-
261 L	4	+++	++	-	-	-
265 L	4	++	++	-	+/-	-
266 R	4	+	++	-	+/-	-
267 R	4	++	+	-	+/-	-
268 L	5	++	+	-	-	-
270 L	5	++	++	-	+/-	-
273 L	5	++	+	-	+/-	-
274 L	6	++	+/-	-	-	-
277 L	4	+	+	-	-	-
278 R	4	+	+	-	+/-	-
279 R	4	+++	+++	-	+/-	-
280 R	4	++	+++	-	+/-	-
283 L	4	++	+++	-	-	-
284 L	8	++++	+++	-	-	-
286 L	6	+	+	-	-	-
NOSEFT02 R	7	+	+	-	+/-	-
NOSEFT05 L	4	+++	+	-	+/-	-
1M R	4	++	++	-	-	-
2M R	3	++	+++	-	+/-	-
Controls	**Passage**	**AE1:AE3**	**CK7**	**CA125**	**FSP**	**F8**
T71	23	++++	+++	++	n/a	-
INOF2	5	-	-	-	++++	n/a
HUVEC	3	-	-	n/a	+/-	+++

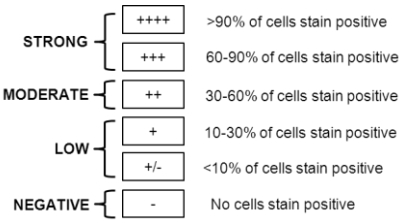


Table 3.1: Summary of fluorescent immunocytochemistry results for the established NOSE primary cell lines. Table presents the scoring of immunostaining for the selected markers AE1:AE3, CK7, CA125, FSP and FVII. The control cell lines used to assay positive and negative control staining for the markers are shown at the bottom of the table. Highlighted in red are cell lines that exhibited a strong positive staining for FSP and were excluded from this study. (n/a= not applicable, where staining was not necessary and was not performed.

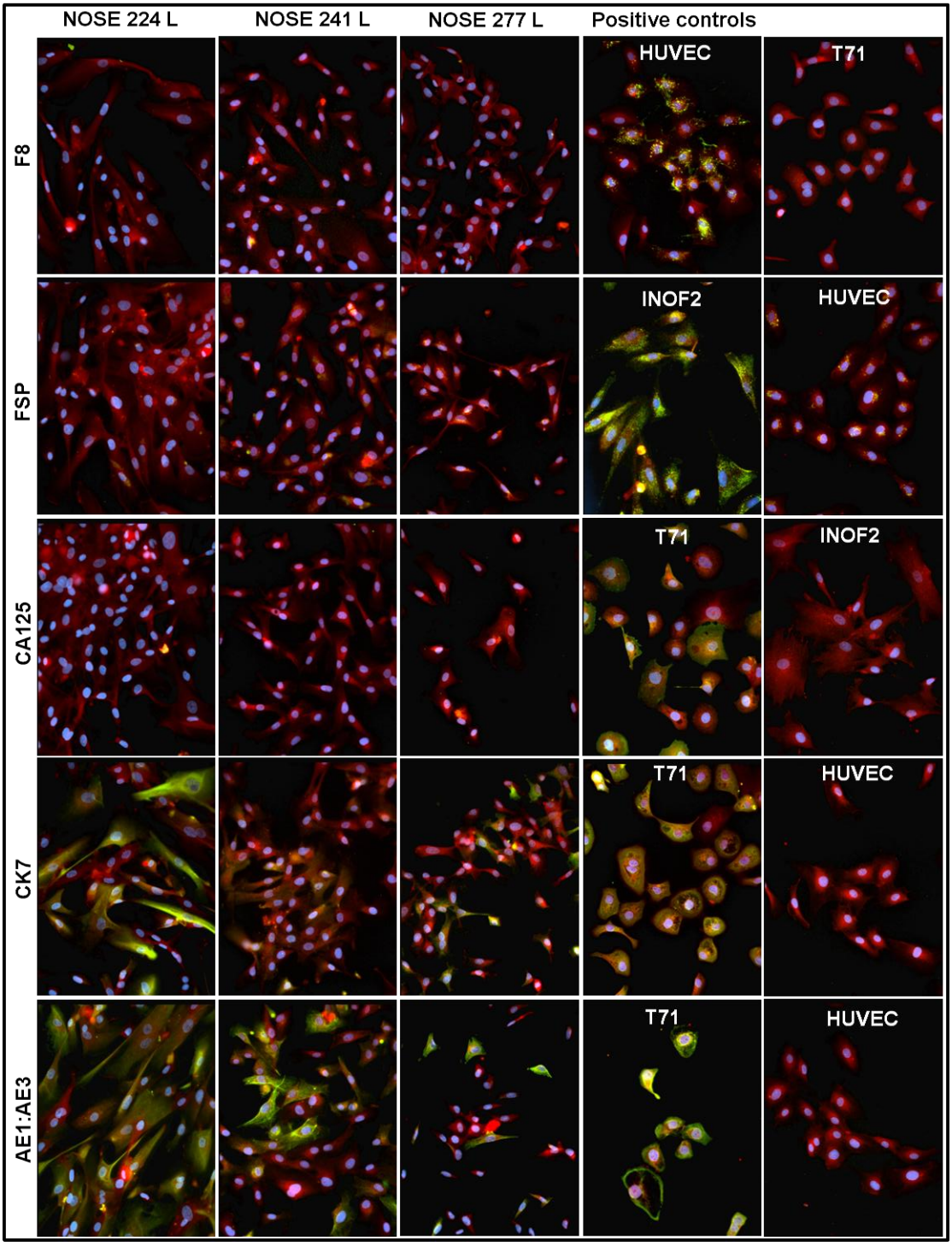


Figure 3.2**:** Examples of fluorescent immunocytochemistry of NOSE cell lines. NOSE cell line expression of AE1:AE3, CK7, CA125, FSP, FVII was captured under a fluorescent microscope (Images at 200×). The blue stain is DAPI that binds to A-T regions of DNA enabling visualisation of the nuclei. The red stain is Evans Blue that is staining the cytoplasm. Finally, the green stain is indicating the expression of the marker of interest. None of the cell lines was found to express FVIII or CA125 with the positive control cell lines HUVEC and T71 respectively showing strong expression. Very low expression of FSP is shown for NOSE 241 L. One example each for strong, mild and low AE1:AE3 epithelial staining is shown for NOSE 224 L, NOSE 241 L and NOSE 277 L respectively.

3.3 Establishing and characterising primary fallopian tube epithelial cell lines

3.3.1 Collection and culturing of primary Fallopian tube epithelial cell lines

Histological and molecular analyses suggest that a proportion of ovarian cancers may be originating from FTE cells. Therefore, it is essential to include FTE cells in any study that investigates EOC. Thus, I aimed to develop primary cultures of FTE cells both in 2D and 3D.

Patients who were due to undergo surgical procedures such as fibroids and polyps resection and total abdominal hysterectomies (TAH) for endometrial cancer were consented for participation on the day of the surgery. The fallopian tubes were initially inspected in surgery by the surgeon in charge as normal and identification of the distal region was made. The fallopian tubes were also later inspected by a histopathologist and confirmed as free of malignancy. This region of the fallopian tube (FT) was dissected and brushings were taken from the lumen of the sub-fimbrial, ampullary region.

The FTE brushings were successfully grown in NOSE-CM media and reached confluency within 10-14 days. In total 8 samples were collected, from patients of ages from 38-65. From these patients, 6 primary cell cultures have been maintained and successfully sub-cultured (Table 3.2), and out of the remaining 2 one senesced very early and one did not grow. In, addition FT tissue was collected from 2 patients and embeded into paraffin as control tissue with wich to compare 2D and 3D cultures.

The morphology of the primary FTE cells *in vitro* was quite heterogeneous containing: a) swirly (endometrial –like cells) b) epithelial cells (small and cuboidal), c) mesenchymal looking cells (larger and elongated) (Figure 3.3). In an attempt to characterise separately the different morphologies observed distinct colonies were initially picked from the different 3 morphologies of the cells. After one passage though they all looked morphologically the same resembling more the mesenchymal morphology initially observed. FTE cells possibly undergo EMT very quickly, but still possessing more epithelial morphology than NOSE, as FTE cells are smaller, with a regular polygonal

morphology and dense cytoplasm. Of the six cell lines obtained the morphology of FT283 and FT284 cell lines resembled more the morphology of the NOSE cell lines being more mesenchymal-like and elongated. The morphology of FT05 was also distinct from the rest with cuboidal and smaller cells that proliferated very aggressively in a similar manner to cancerous cell lines.

Cell line	Patient age	Clinical Data	Histopathological diagnosis	Specimen Collected	Menopause status	Histopathology reports
FT01	51	Mucinous cyst on the left ovary	Benign follicular and epithelial inclusion cysts. Benign serous cystadenofibroma.	Lumen brushing, ampullary brushing, fimbrial end brushing, minced tissue.	No information	Fallopian tube normal
FT02	50	Benign fibrosis	Leiomyomas and folicular ovarian cysts	Lumen brushing	No information	Mild chronic salpingitis But no evidence of malignancy
FT03	38	Fibroid. Adenomyosis	Secretory endometrium and benign leiomyoma	Lumen brushing, TISSUE in formalin	Premenopausal	Fallopian tube normal
FT05	54	Complex hyperplasia with atypia	Autolysed endometrium. No evidence of invasive carcinoma. Benign leiomyomas.	Lumen brushing	Postmenopausal	Both onaries with post men features and fallopian tubes normal
FT283	65	Grade 2-3 endometrial cancer	Grade 2 endometrioid adenocarcinoma with more than 50% myometrial invasion	Lumen brushing	Postmenopausal	Ovaries post menopausal features. Fallopian tube normal
FT284	60	History proven endometrial cancer	Grade 2 endometrioid adenocarcinoma	Lumen brushing	Postmenopausal	Fallopian tube normal

Table 3.2: Information regarding the established primary FTE cell lines. Tabulated are patient age and diagnosis, menopausal status and histopathological reports for the specimens.

Using the protocol described for collection and culture of FTE cells, I have successfully sub-cultured primary FTE cells for over 13 passages. The growth curves for FTE cell lines FT01, FT02, FT03, FT05, FT283 and FT284 are shown on Figure 3.5. The growth curves were started at passage 3 to passage 5 for each cell line and were performed in triplicate. After 11-13 passages 100% of the cells underwent senescence, as assayed by senescence-associated β-galactosidase expression (Figure 3.4). The lifespan of FTE cells has been <60 days similarly to previous studies reporting the lifespan of NOSE cell lines (Lawrenson et al, 2009). The average population doubling was calculated for each of the cell lines and ranged from 16 to 21. The mean time for a population doubling to occur ranged from 21 to 42 hours (Table 3.3).

The FTE cell line FT05 was behaving differently than the other 5 in culture as it was sub-cultured for more than 18 passages in contrast with the rest of the cell lines that possessed a similar lifespan to the NOSE cell lines of maximum 12 to 13 passages.

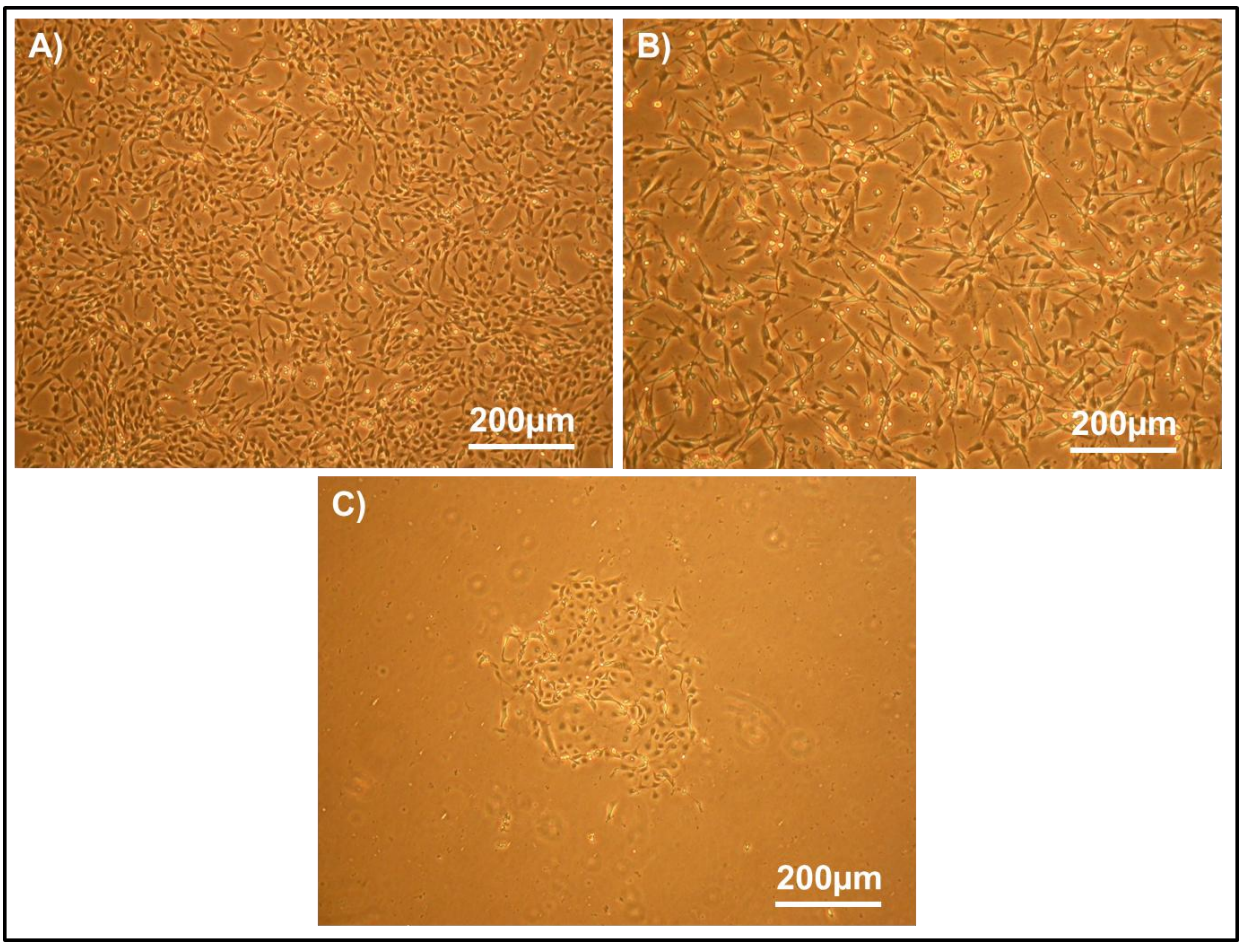


Figure 3.3: Different morphologies of primary FTE cultures. A) epithelial-like B) mesenchymal like, C) endometrial-like.

Days in culture	Population Doublings (PD)					
	FT01 PD	**FT02 PD**	**FT03 PD**	**FT05 PD**	**FT283 PD**	**FT284 PD**
0	0.0	0.0	0.0	0.0	0.0	0.0
3	3.1	2.0				
4				4.0	3.9	2.0
5	6.1		3.4			
7				7.7		
8		5.4			6.8	
9	11.0		6.8	10.1		
10						5.7
11				12.2		
12	14.1	7.4			9.6	
13			9.4			7.4
14					11.9	
15	16.7			16.6		
18	19.3	9.8				10.2
19			11.6	19.1	14.5	
22					17.8	
23	22.4	13.9		22.7		12.5
27			13.2			
28	24.4	16.0			21.2	14.3
29				26.0		
33			15.2	26.9		
34		16.6			24.2	15.3
36				30.1		
39				31.8		
40	26.3	19.7	15.9			
41					26.6	
44		22.4				
46				32.8		
49					29.2	
51		23.2				
55					32.4	
60				35.4		
Average PD	**15.9**	**13.6**	**9.6**	**10.8**	**18.0**	**21.2**
Mean PD time (hrs)	27.4	29.2	37.2	38.2	21.6	42.1
St Dev of PD time (hrs)	1.4	0.3	1	2	0.5	0.8

Table 3.3**:** Population doublings of the established primary FTE cell lines. The average population doublings time (hours) was calculated based on the median population doubling for each cell line.

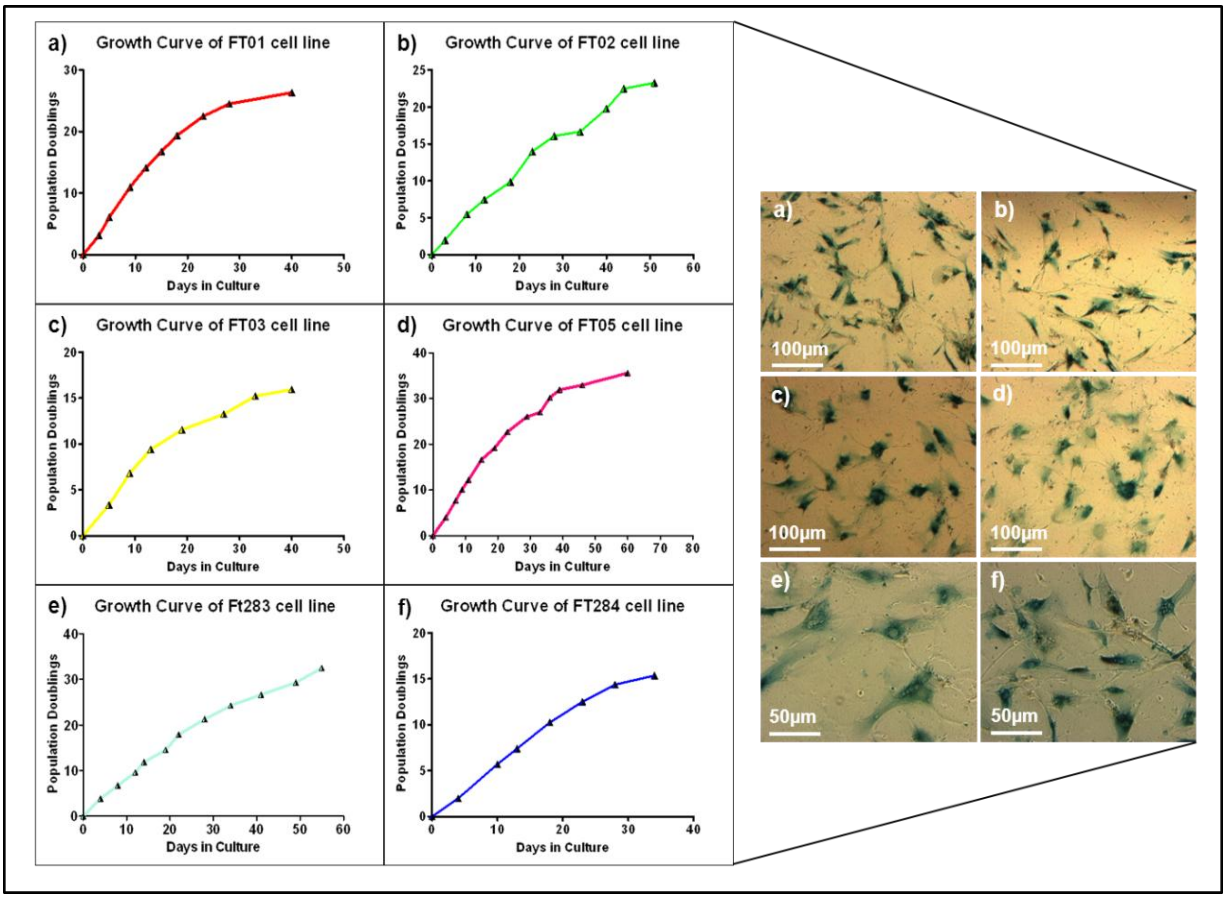


Figure 3.4**:** Growth curves and β-galactosidase staining for senescence of FTE cell lines. a) FT01, b) FT02, c) FT03, d) FT05, e) FT283, f) FT284. The lifespan of the cells did not exceed 60 days after which point all the cell lines reached senescence as shown by staining for β-galactosidase expression (blue colour).

3.3.2 Karyotyping of established FTE cell lines

Cells of normal origin, that are not transformed should possess a normal karyotype, therefore the karyotype of the 6 established FTE cell lines was analysed. Cells of low passage number were sparsely seeded in a 25cm^2 flask a day before they were collected and chromosomal spreads were prepared. The karyotypes were then examined by a qualified clinical cytogenetistist under a light microscope (TDL genetics, London). The number of chromosomes as well as their length, the position of the centromeres, banding pattern and any other physical characteristics was commented on to give a detailed description of any abnormalities. Five out of 6 cell lines had a normal model karyotype of 46, XX with two of them having a small proportion of cells (3/20 and 1/20) carrying a structural abnormality, possibly a harvesting artefact as commented by the cytogenetisist. One out of the six cell lines was found to have karyotypically abnormal cells in 20/20 cells that were analysed. This cell line was FT05, and its abnormal karyotype could explain the aggressive proliferative phenotype this cell line shows since it is clearly a transformed cell line. The karyotypes of the FTE cell lines are shown in Figure 3.17 and a summary of the karyotype descriptions is provided in Table 3.4.

Cell line	Passage	Total number of cells analysed	Karyotype	Karyotype description
FT01	p4	20	46, XX	Model Karyotype showed a normal female chromosome complement and banding pattern.
FT02	p5	17	46, XX, [17]/47, XX, +2 [2]/46,XX, DER,(11) t (11; acro p arm) (p15; p11.2) [1]	Model karyotype showed a normal female chromosome complement and banding pattern. However 2 cells were found with trisomy 2 and 1 cell seemed to have a structurally abnormal chromosome 11 with the satelite region of an acrocentric chromosome attached to its short arm at break point p15.
FT03	p3	20	46, XX	Karyotype of 20 cells showed a normal female chromosome complement and banding pattern
FT05	p3	20	46,XX,t(6;6)(p21;q21) [11]/ 47,XX, t (6;6) (p21;21), +mar [2]/46,XX, t (2;3) (p21; q13) [1]	No karyotypically normal cells were found
FT283	p7	10	46, XX	Model Karyotype showed a normal female chromosome complement and banding pattern. A single cell with structurally abnormal karyotype was seen.
FT284	p4	6	46, XX	A model Karyotype showed a normal female chromosome complement and banding pattern.

Table 3.4: Karyotype description of the established primary FTE cell lines. Karyotyping analysis revealed a normal model karyotype for 5 of the 6 cell lines collected.

Figure 3.5**:** Karyotypes of the established FTE cell lines. The karyotyping revealed a normal model karyotype for 5 of the 6 cell lines collected. Cell lines FT01, FT02, FT03, FT283 and FT284 have a karyotype of a normal female chromosome complement and banding pattern (46, XX). Cell line FT05 however, revealed a female karyotype with a complement of 46 chromosomes and an apparently balanced reciprocal translocation between one chromosome 6 at breakpoint p21 and its homologue at breakpoint q21 as indicated by the arrows in the figure above.

3.3.3 Immunofluorescent staining of FTE cell lines

I chose to characterise the fallopian tube cell lines established by staining with a panel of markers to establish their epithelial origin and identify which epithelial markers they mostly express and to confirm that they are not contaminated with other cell types during collection.

AE1:AE3 **:** Pan-cytokeratin. This marker was used in order to confirm the epithelial status of the FTE cell lines.

CK7 **:** Cytokeratin-7. This marker was also used in order to confirm the epithelial status of the FTE cell lines.

BerEP4 **:** Ep-Cam. Ep-Cam is an epithelial specific antigen. It is a glycoprotein located on the cell surface of epithelial cells and has been linked to epithelial differentiation (Latza et al, 1990) and found to be expressed in many epithelial carcinomas. It is also used a marker for cells of epithelial origin but expression in high levels may indicate neoplastic transformation.

CA125 **:** Cancer antigen 125. Although this is a marker that indicates malignancy in NOSE cell lines, it is also a lineage marker that is expressed by Mullerian epithelium during Mullerian development and its expression is retained in FT tissues (Auersperg, 2001). This marker was used to investigate whether the FTE cell lines would retain Mullerian lineage specific markers when isolated and cultured.

FSP **:** Fibroblast specific protein. This marker was used in order to confirm that the FTE cell lines did not have fibroblast contamination from the stromal cells underlying the mucosa of the tubes. Having seen morphological signs of EMT *in vitro* while sub-culturing the FTE cell lines, low levels of FSP expression were expected to be observed.

Laminin **:** Laminins are glycoproteins that are forming the extracellular matrix. They are structural support proteins in the microenvironment of epithelial tissue holding the cells together. Laminin has been shown to play a role in the development of some types of cancer such as colon cancer (Kitayama et al, 1999) and breast cancer (Beliveau et al, 2010) and even ovarian cancer (Poon et al, 2011)

Vimentin : Vimentin is a type II intermediate filament protein. Vimentin is part of the cytoskeleton of cells together with actin and tubulin and is strongly expressed by mesothelial cells, and in cells that undergo epithelial-to-mesenchymal transition. Vimentin has been shown to be highly expressed in different types of carcinomas such a colon and breast (McInroy and Maatta, 2007). It has also been shown that Vimentin expression is inversely associated with keratin expression in breast cancer (Thomas et al, 1999). The same study showed that relative keratin and Vimentin expression can better indicate the prognosis and tumour phenotype than investigating the two markers independently.

E-Cadherin : E-Cadherin is a transmembrane protein with a role in cell adhesion and is very important for the architectural organisation of tissues. E-cadherin is not expressed in ovarian surface epithelium (OSE) but it is expressed in the epithelium of the inclusion cysts and it is also expressed in epithelial ovarian carcinomas. Screening FTE primary cell cultures for expression of this marker will provide an insight for their proposed role as cell of origin for EOC.

FVIII : Factor VIII antibody. This marker was used to confirm that there was no blood cell contamination in the FTE cell lines.

Each immunofluorescence staining experiment for characterisation of the FTE cell lines was performed with positive and negative controls (Figure 3.6) similarly to the NOSE cell line characterisation for markers, AE1: AE3, CK7, CA125, FSP and FVIII. T71was also used as a positive control for marker BerEP4. INOF2 was also used as positive control for Vimentin and laminin expression. FTE cell lines FT01, FT02, FT03, FT05, FT283 and FT284 were characterised for the markers described between passages 5-7 and pictures were obtained under the 200x of a fluorescent microscope (Figures 3.7, 3.8 and 3.9). A summary of the staining is shown on Table 3.5.

Cell line	Passage	AE1:AE3	CK7	BerEP40	FSP	CA125	Laminin	Vimentin	E-Cadherin	FVIII
FT01	7	-	++++	-	-	-	+++	++	-	-
FT02	7	+++	+	-	-	+++	+++	++++	-	-
FT03	7	++	+/-	-	-	-	++++	++++	-	-
FT05	5	-	-	-	+/-	-	+++	+++	-	-
FT283	7	++	+	-	-	-	++	++	-	-
FT284	5	+++	++	+/-	-	-	+++	++++	-	-

Table 3.5: Summary of i fluorescent immunocytochemistry results for the established FTE cell lines. The intensity of the staining is presented with a range from (-) to (++++) similarly to Table 3.2 for the NOSE cell line staining. Interestingly, only one of the FTE cell lines is expressing the marker CA125 and none of them are expressing E-Cadherin. Vimentin is also moderately to strongly expressed in all the cell lines which is an indication of the EMT they are undergoing noticeable morphologically as the passage number increased while culturing them.

All the cell lines were free of cell contaminants such as stromal cells and endothelial cells since they stained negative for the markers FSP and FVIII respectively. Low FSP expression was observed in FT05 but this could be attributed to EMT the cells undergo rather than stromal contamination, especially after comparing with the high expression of FSP by the stromal line INOF2 shown in Figure 3.6.

Most of the cell lines had in common the fact that they were negatively stained for marker CA125 apart from cell line FT02. None of the FTE cell lines were found to express E-Cadherin or BerEP4 which are found expressed in EOCs. Previous work in the lab with the NOSE cells has shown that they do not express BerEP4 either in similar *in vitro* cultures (Kate Lawrenson, personal communication). Additionally, all of the cells lines were found to stain positive for Vimentin and laminin.

A striking observation regarding cell line FT01 was the absence of AE1:AE3 and strong positive expression of CK7. This was in contrast with the NOSE cell lines and also primary FTE cell lines FT01, FT02, FT283 and FT284 in which AE1:AE3 was the epithelial marker found to be expressed in higher levels than CK7. FT05 stained negative for both AE1:AE3 and CK7 epithelial markers. The abnormal staining results may be related to the chromosomal abnormality detected in this cell line. These data may suggest that FT05 has undergone some degree of neoplastic transformation. More work would be needed to fully characterize the malignant phenotype of this cell line.

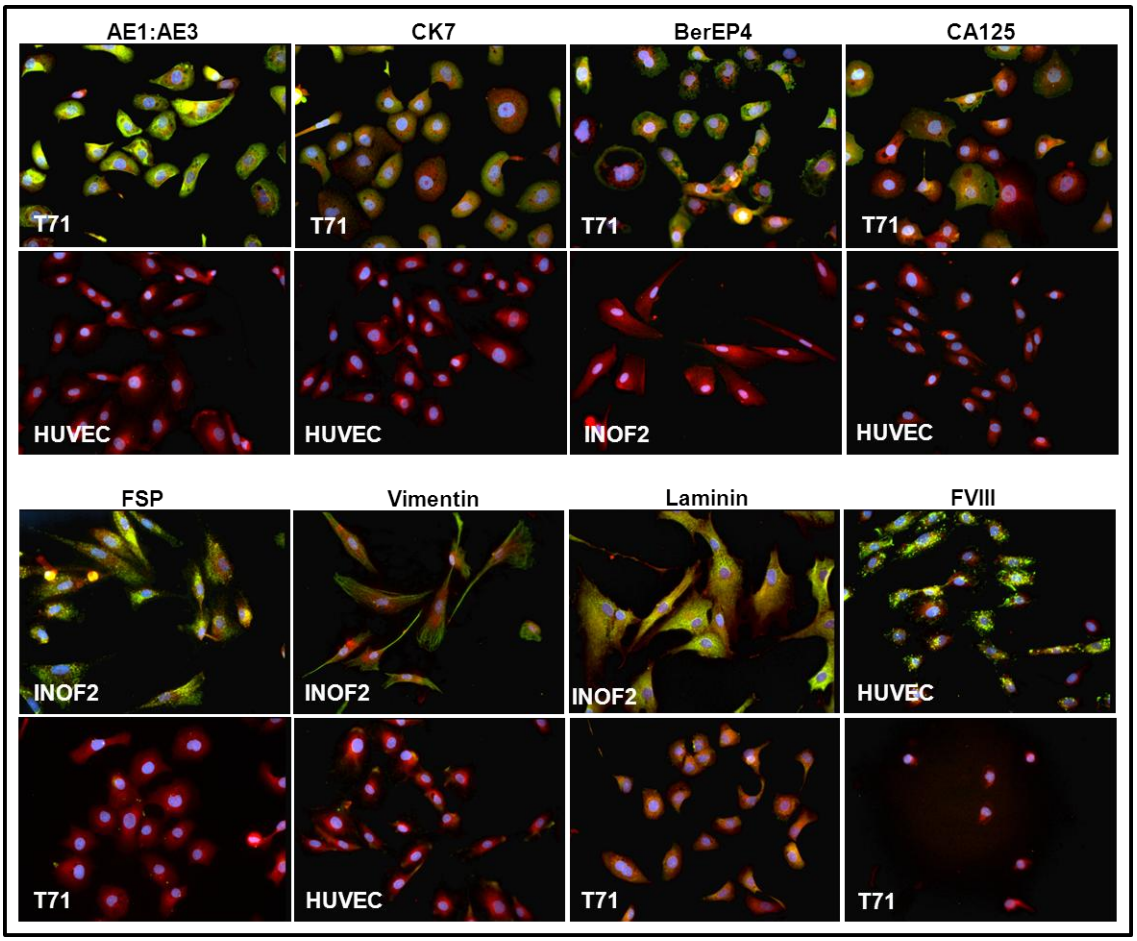


Figure 3.6**:** Fluorescent immunocytochemistry of positive and negative control cell lines for nine markers. Images were captured under a fluorescent microscope (Images at 200×). The blue stain is DAPI enabling visualisation of the nuclei. The red stain is staining the cytoplasm. Finally, the green stain is indicating expression of the marker of interest.

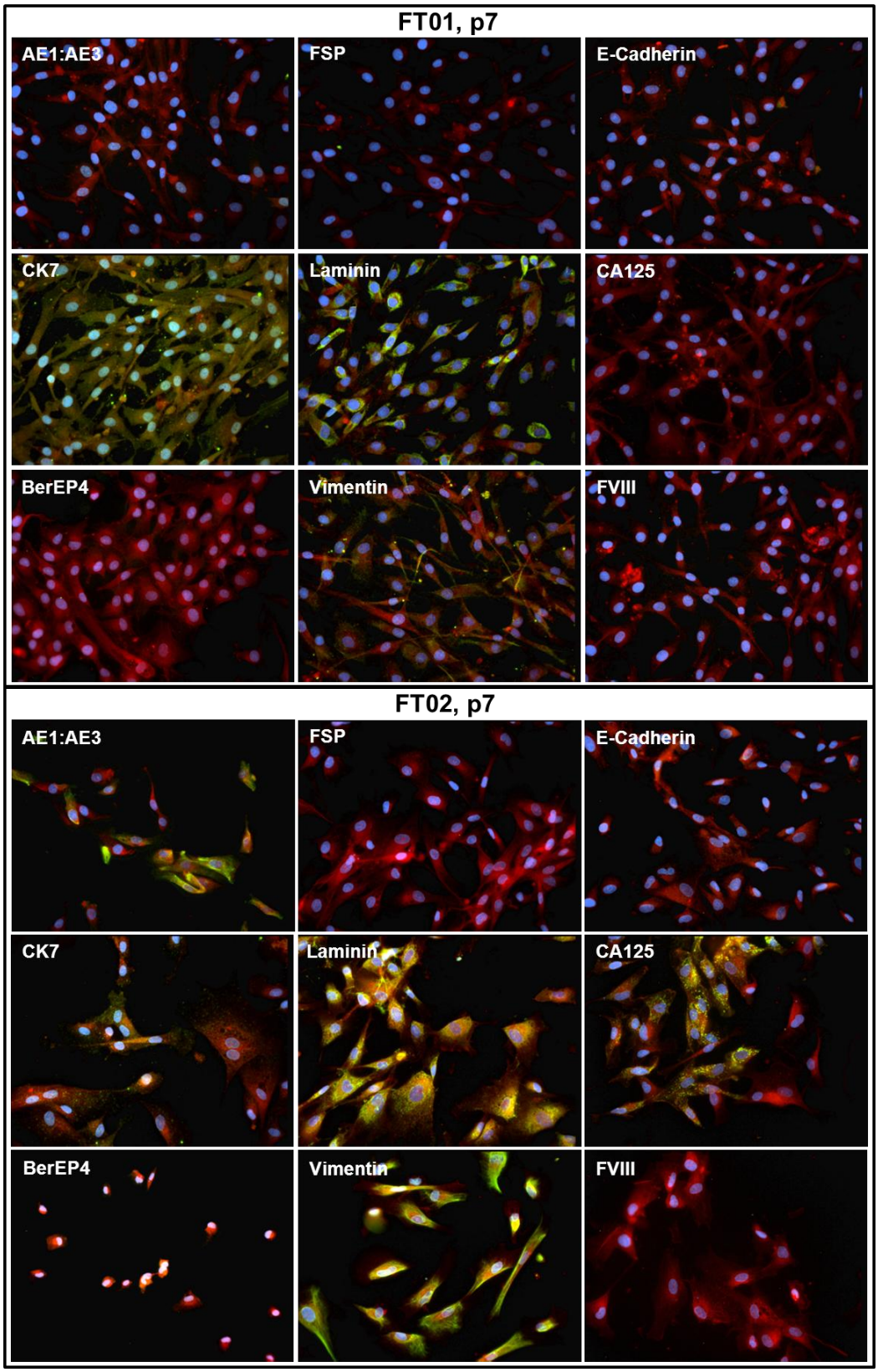


Figure 3.7: Fluorescent immunocytochemistry of FT01 and FT02 primary cell lines for nine markers (Images at 200×). The blue stain is DAPI. The red stain is Evans Blue. The green stain is indicating expression of the marker of interest. The positive and negative controls for these markers are shown in figure 3.6.

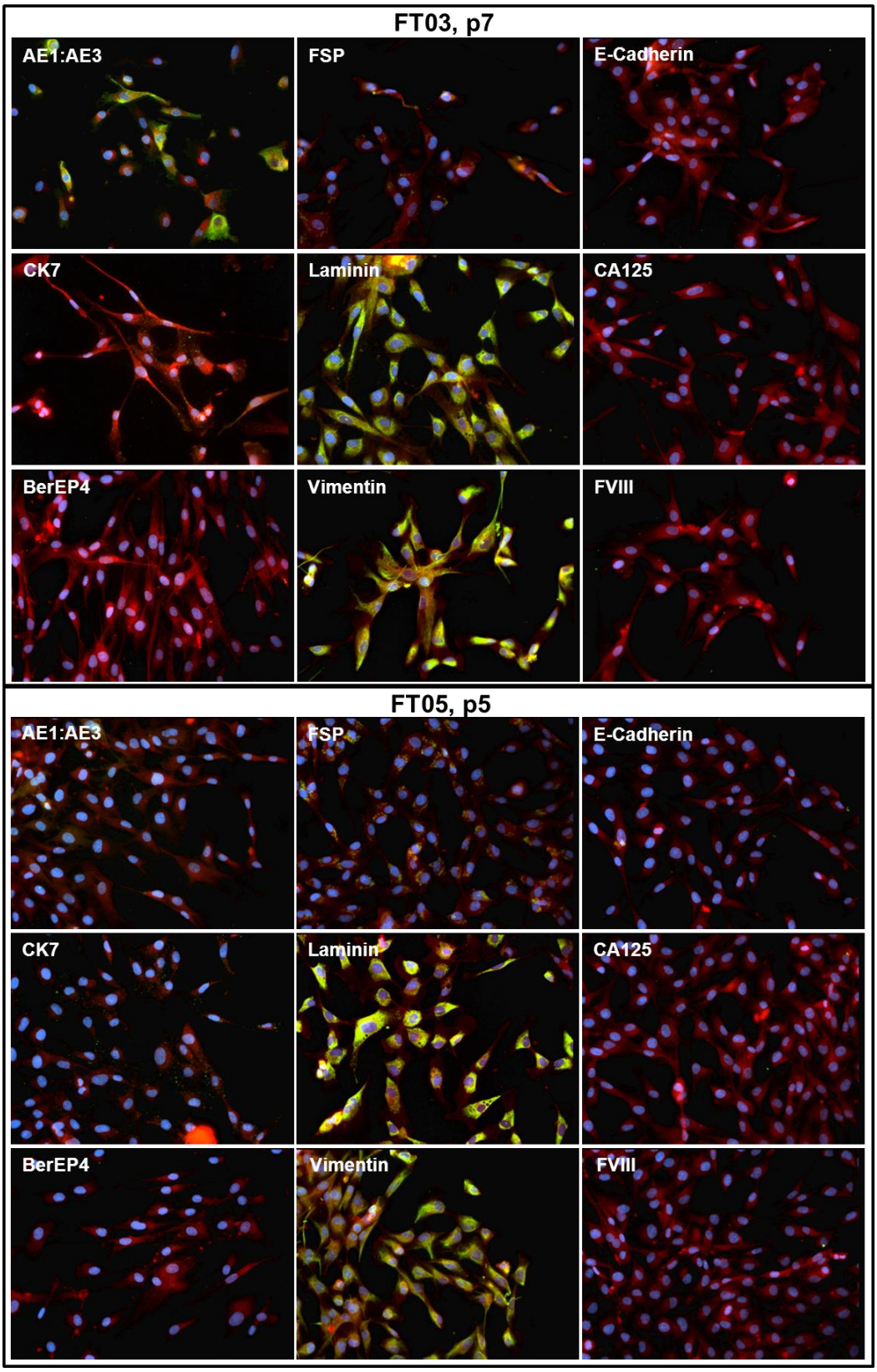


Figure 3.8: Fluorescent immunocytochemistry of FT03 and FT05 primary cell line for nine markers (Images at 200×). The blue stain is DAPI corresponding to nuclei. The red stain is staining the cytoplasm. Finally, the green stain is indicating expression of the marker of interest. The positive and negative controls for these markers are shown in figure 3.6.

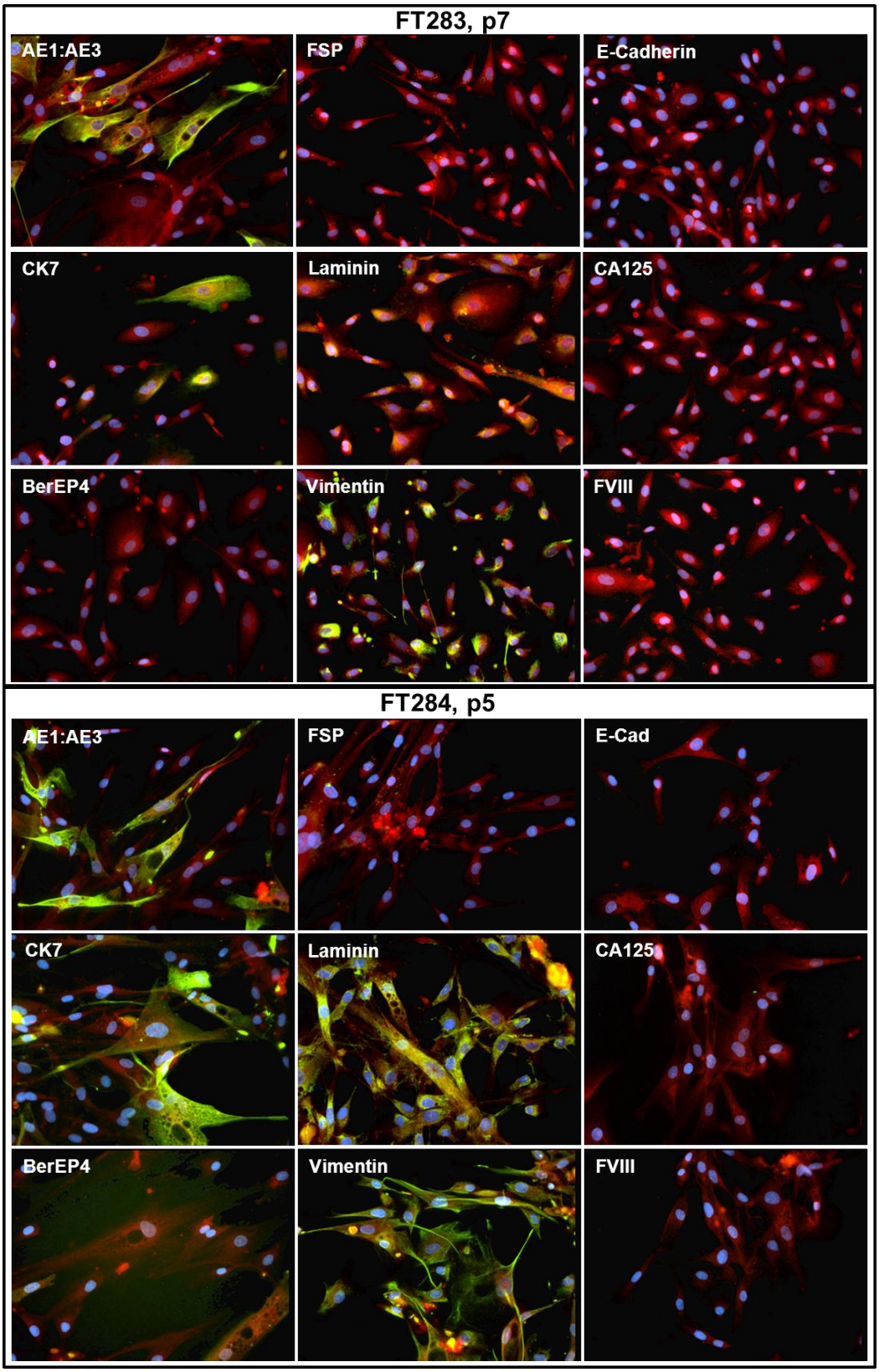


Figure 3.9: Fluorescent immunocytochemistry of FT283 and FT284 primary cell lines for nine markers (Images at 200×). The blue stain is DAPI corresponding to nuclei. The red stain is staining the cytoplasm. Finally, the green stain is indicating expression of the marker of interest. The positive and negative controls for these markers are shown in figure 3.6.

3.3.4 Isolated FTE cultures are rich in secretory cells confirmed by lineage marker PAX-8

In order to establish an *in vitro* model for studying EOC carcinogenesis from FTE cell lines it was important to ensure that they recapitulate the histology of the FT tissue and will express lineage specific markers. Fallopian tube epithelial mucosa cells *in vivo* are separated into secretory and ciliated cells. When the FTE cells were cultured the cilia were lost either due to culturing or possibly revertion of the cells to a secretory type phenotype. The proposed cell of origin for many serous EOC is the secretory cell of FT epithelium (FTSEC) (Levanon et al, 2008, Kurman et al, 2011). Previous work performed in an *ex-vivo* FT system has established that an FT secretory cell lineage marker is PAX8 (Levanon et al, 2010). PAX8 is a paired box nuclear protein with a role as a transcription factor that triggers expression of genes involved in thyroid and reproductive tract development. Staining of the FTE cultures with PAX8 aimed to elucidate whether they express this lineage specific marker and the percentage of the expression would show whether the secretory cells predominantly maintained their characteristic in these cultures. In Figure 3.11 it is shown that all the FTE cell lines established expressed PAX8 localised as expected in the nucleus. The expression of PAX8 by 100% of FTE cells indicates that the ciliated FTE cells when cultured may be reverting to a secretory phenotype. Each PAX8 immunofluorescence staining experiment for characterisation of the FTE cell lines was backed by positive and negative controls (Figure 3.10). T71, an endometrial cancer cell line, was used as a positive control which demonstrates 100% strong expression of Pax-8.

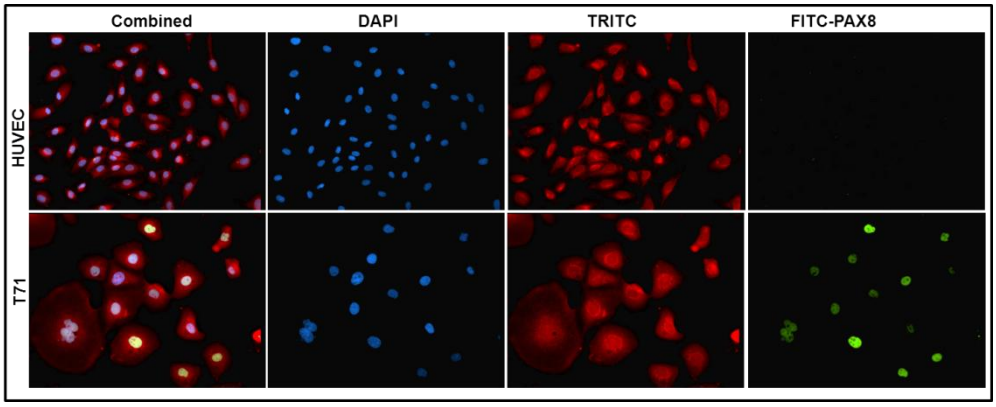


Figure 3.10**:** Fluorescent immunocytochemistry of positive and negative control cell lines for PAX8 marker. Images were captured under a fluorescent microscope (Images at 200×).

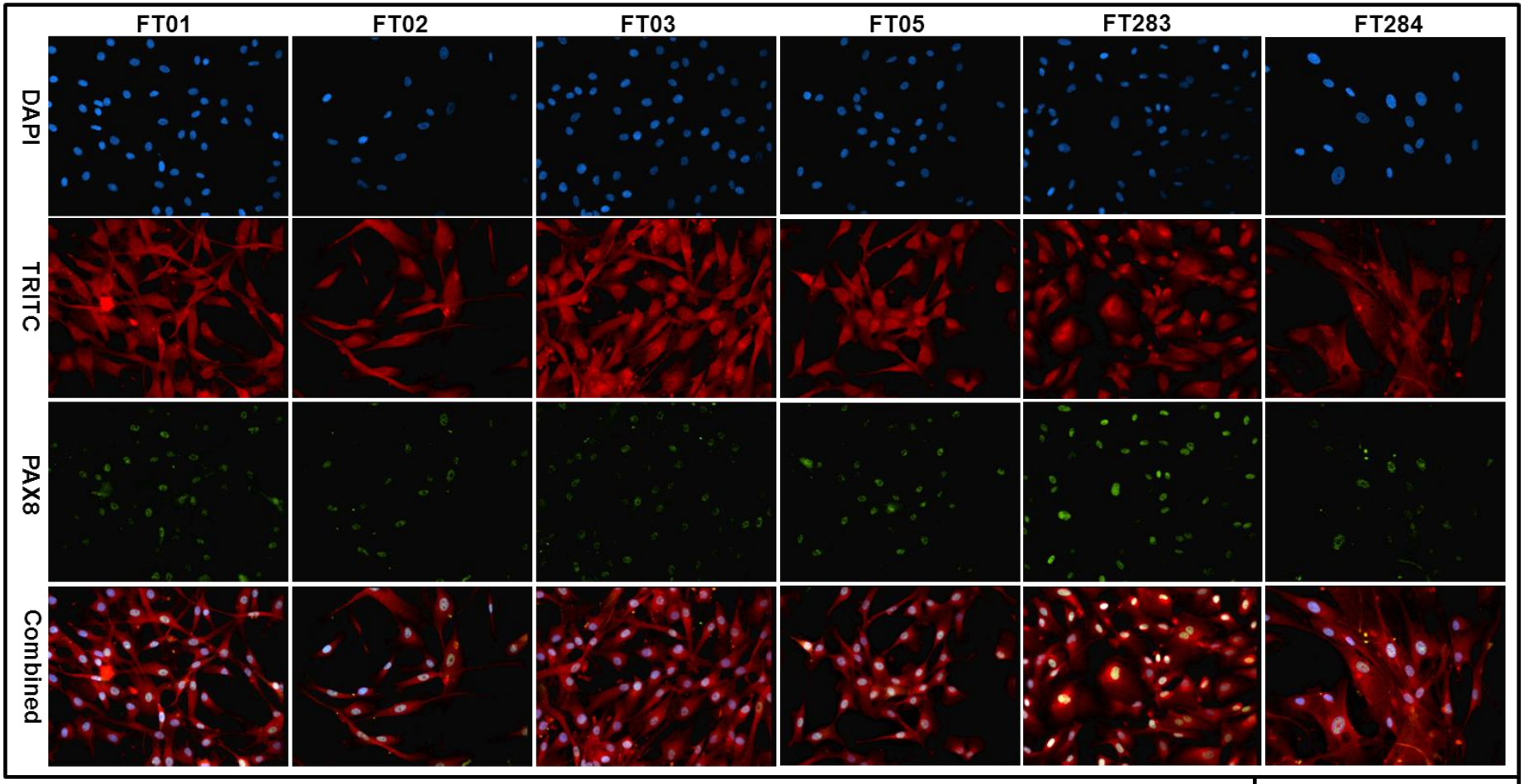
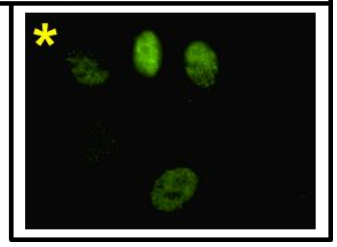


Figure 3.11**:** Fluorescent immunocytochemistry of FTE cell lines for lineage marker PAX-8 (Images at 200×).The blue stain is DAPI, enabling the visualisation of the nuclei. The red stain stains the cytoplasm. Finally, the green stain is indicating expression of PAX8 which is expressed by 90-100% of the cells in culture. **(*)** Nuclear localisation in a punctate pattern is observed as shown at the 1000× magnification image. The positive and negative controls for for PAX8 are shown in figure 3.10.



3.3.5 Investigating the efficiency of FTE cell lines to grow in an anchorage independent manner

Transformed cell lines have the ability to grow in an anchorage independent manner in soft agar. However, non-malignant cells should not have the capacity to grow colonies in soft agar. The efficiency of FTE cell lines to grow in an anchorage independent manner was assayed by performing soft agar assays with different number of initially plated cells (20000, 10000 and 5000 cells) for 4 weeks. All assays were performed in triplicate and results were confirmed in independent experiments. Interestingly and unexpectedly, all the FTE cell lines that were subjected to anchorage independent growth assays formed colonies even when seeded in low density (5000 cells/well). The colony forming efficiency (CFE) was calculated for all the cell lines and is shown on Figure 3.12. A detailed table with the colony counts for each replicate is shown in Appendix 1, Table 2. The CFE of FT05 which was a chromosomally abnormal, possibly transformed, cell line was more aggressive closely resembling the efficiency of TOV112D EOC cell line to grow in an anchorage independent manner. However, the colonies formed were more similar in size to the colonies of the other FTE cell lines.

The morphologies of the FTE colonies in soft agar were closely inspected and did not resemble the morphology observed in the cancer cell line TOV112D with the exception of FT05 which was more compact and dense than the rest (Figure 3.13). TOV112D colonies formed in soft agar were much larger and compact but for the FTE cell lines the colonies formed were much smaller and possessed a ring, looking like cells migrating to or from the colony.

In order to achieve a better visualisation of the structure of the FTE colonies formed in soft agar, the colonies were embedded in paraffin, cross sectioned and stained with Heamatoxylin and Eosin (H&E) (Figure 3.14). Observation of the stained cross sections clarified that the ring observed on the outside of the colony was not formed by cells migrating from or towards the colony but rather debris from cells as they were under sub- optimal growing conditions showing karyorexis and producing cellular debris. The soft agar FTE colonies were very small ranging between 20-100μm in size compared to the >500μm colonies of EOC cell line TOV112D. Cross sections of various sizes of

these colonies revealed that instead of increased proliferating ability and efficiency to grow in soft agar in the aggressive way observed in cancer cell lines, the FTE colonies resembled more closely cell aggregates. Most colonies and particularly slightly bigger colonies such as the ones formed by FT03 were seen to produce a hyline type material typically observed in 3D cultures (Figure 3.14).

Number of cells seeded	FT01, p8	FT02, p8	FT03, p5	FT05, p5	FT283, p8	FT284, p8	112D
1000							59
5000	11	7	13	35	8	12	
10000	10	7	14	37	6	11	
20000	11	6	12	34	4	11	
CFE average (%)	**11**	**7**	**13**	**36**	**6**	**12**	**59**

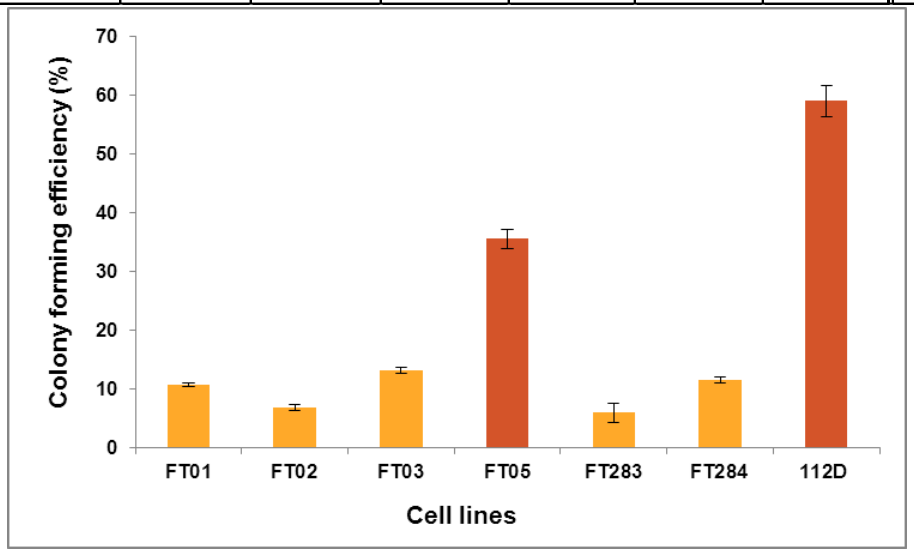


Figure 3.12**:** Anchorage independent growth efficiency for FTE cell lines. The colonies were counted after seeding 5000, 10000 and 20000 cells and cultured for four weeks. All FTE cell lines apart from FT05, which demonstrated a very high CFE, had a CFE between 6-13%. TOVTOV112D EOC cell line was used as a positive control for anchorage independent growth. The Table is showing the averaged CFE in soft agars for each of the FTE cell lines. The histogram represents the averaged CFE for each of the tested cell lines.

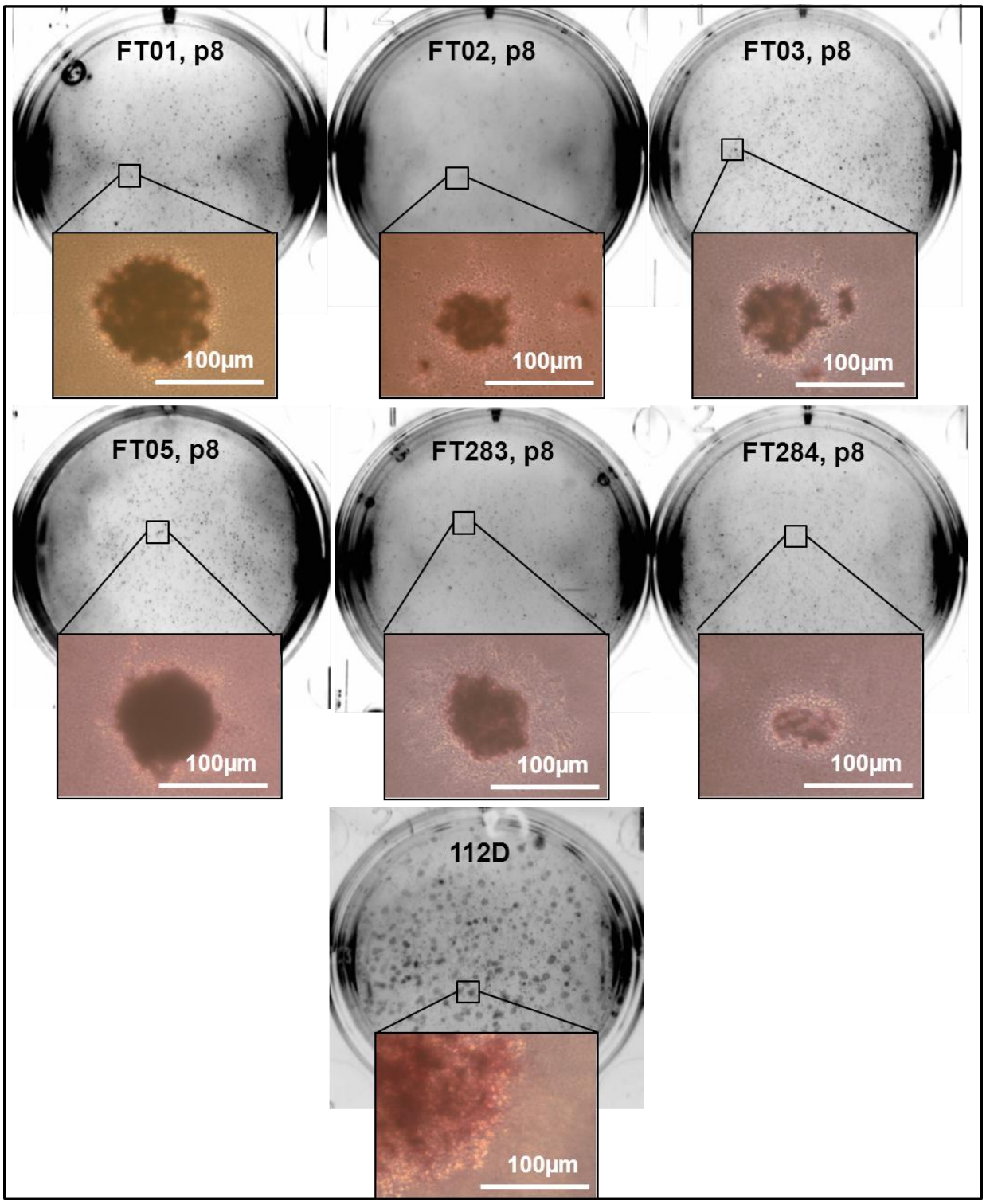


Figure 3.13**:** Colony morphology of FTE cell lines growing in soft agar after seeding 20000 cells and cultured for four weeks. Visibly is the ring of single cells extending form the colonies of the FTE cell lines compared to the TOV112D cell line colonies that appear more compact and with a defined perimeter. Closer investigation shown in figure 3.14 revealed that the rings are most possibly cell debris rather than single cells migrating towards the colonies.

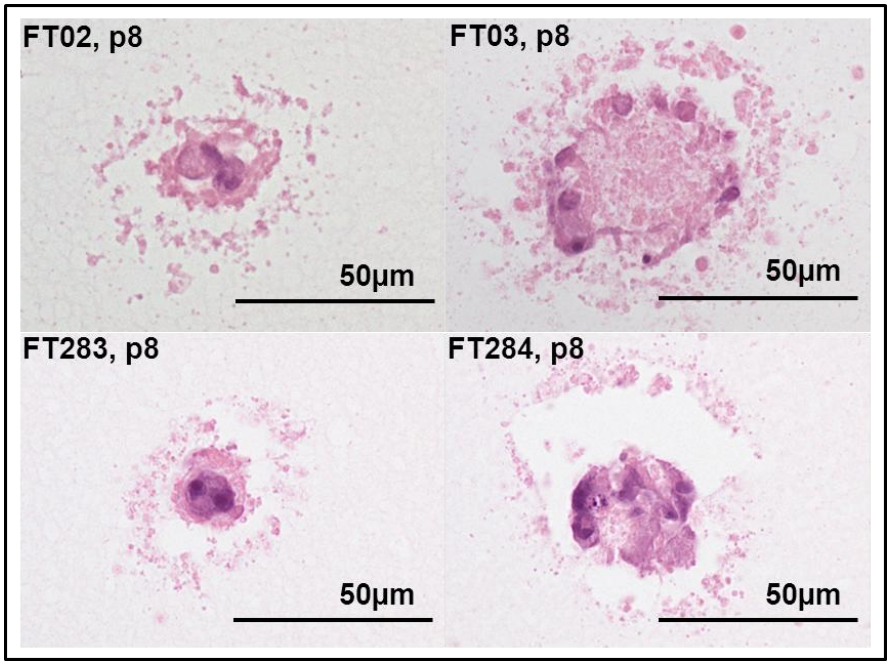


Figure 3.14**:** Representative H&E staining of FTE colonies grown in soft agar. The outer ring seems to be formed by debris produced by the cells during karyorexis.

3.4 Establishing primary FTE cell lines as three-dimensional (3D) spheroid cultures

Studying cells using two-dimensional (2D) *in vitro* models can provide information about the markers they express and the cell line's proliferating abilities but it is limiting as the cell-cell interactions and cell-extracellular matrix contacts that are characterising the *in vivo* tissue are not represented. Creating three-dimensional *in vitro* models offers insight into the microenvironment of the cells more closely resembling the complex tissue architecture. I have established FTE 3D cultures on polyHEMA coated tissue culture plates in an attempt to create a model that recapitulates FT tissue morphology and function in regards with architecture, marker expression and proliferating abilities. FTE cells successfully formed spheroids within 24 hours after seeding and were incubated for 15 days. During this period the 3D structures increased in size and became more compact (Figure 3.15). The spheroids were embedded into paraffin blocks and sections were cut and stained with H&E to enable the study of the internal architecture and histology of the spheroids (Figure 3.16).

Cross sections of the 3D spheroids revealed a similar architecture for 3D cultures of cell lines FT01, FT02, FT03, FT283, and FT284 (Figure 3.16). The spheroids were of varying size and shape. For most of the spheroids a single layer of epithelial like flattened cells was present in most, however multiple layers of peripheral cells were observed in some of them. These circumferential layers of cells in the periphery of the spheroids were mostly consisting of spindle shaped cells. The spheroids had a crescentic cap which was more prominent in cell lines FT01, FT02, FT283 and FT284. The central portion of the spheroids had hyline material with some cellularity which resembles the extracellular matrix material present in the *in vivo* tissue. Some spheroids showed more than one hyline cores (foci) surrounded by spindles of rounded cells with prominent nucleoli. The hyline produced contained a number of viable cells mixed with abundant karyorectic debri (nuclear dust). The viable cells in the inner core of the matrix of the structures exhibited cellular pleiomorphism as some cells had significantly larger nuclei and abundant cytoplasm and some had indistinct cytoplasmic borders. Nuclei pleiomorphism was also obvious with most of cells possessing round nuclei with prominent nucleoli. A few cells

appeared to have spindle shaped nuclei and were more elongated. Within the hyline material there were residual cells that often showed apoptosis and degenerative changes such a karyorexis and cytoplasmic vacuolation responsible for the nuclear dust mixed with the hyline material. Maybe cells produce matrix and then undergo degeneration changes which is why karyorexis was observed throughout the structures. Interestingly, the appearance of hyaline material was sometimes enveloped in cell cytoplasm. This is observed in tumours sometimes. The 3D structures of cell line FT05 had a very different architecture than the rest. Nuclei were more proliferative with oval or round shapes. Spheroids were multi-layered and more cellular with smaller amounts of hyline matrix. Cap structures were not as prominent as in the other cell lines. Spheroids were very cellular in central matrix with a lot of karyorexis probably corresponding to increased cellularity. Some spheroids had small hyline bodies and hyline material seemed to form where cellular growth was smaller.

3D cultures were created for 15, 26 and 40 day incubation periods (Figure 3.15) in order to see whether increased apoptosis within the spheroids was an attempt of the 3D structure to recapitulate the FT architecture *in vivo* and a lumen would be formed. 3D cultures incubated for longer than 15 days displayed an increased proportion of apoptotic cells. The increased rate in apoptosis and karyorexis was subtle and may have been related to the limited availability of nutrients for the inner core of the spheroids as the structures increased in size.

Overall, comparing the 3D cultures of FTE and NOSE I found that FTE spheroids also presented crescentic caps similarly to the NOSE however less prominent that the NOSE. The presence of hyline material was found in both NOSE and FTE 3D structures but FTE spheroids appeared to have more cellularity within the cores (Figure 3.16). Similar to NOSE, the architecture observed in FTE 3D structures more closely resembles the *in vivo* FT architecture of cell interaction and interaction of cell with extracellular matrix than the 2D FTE cultures.

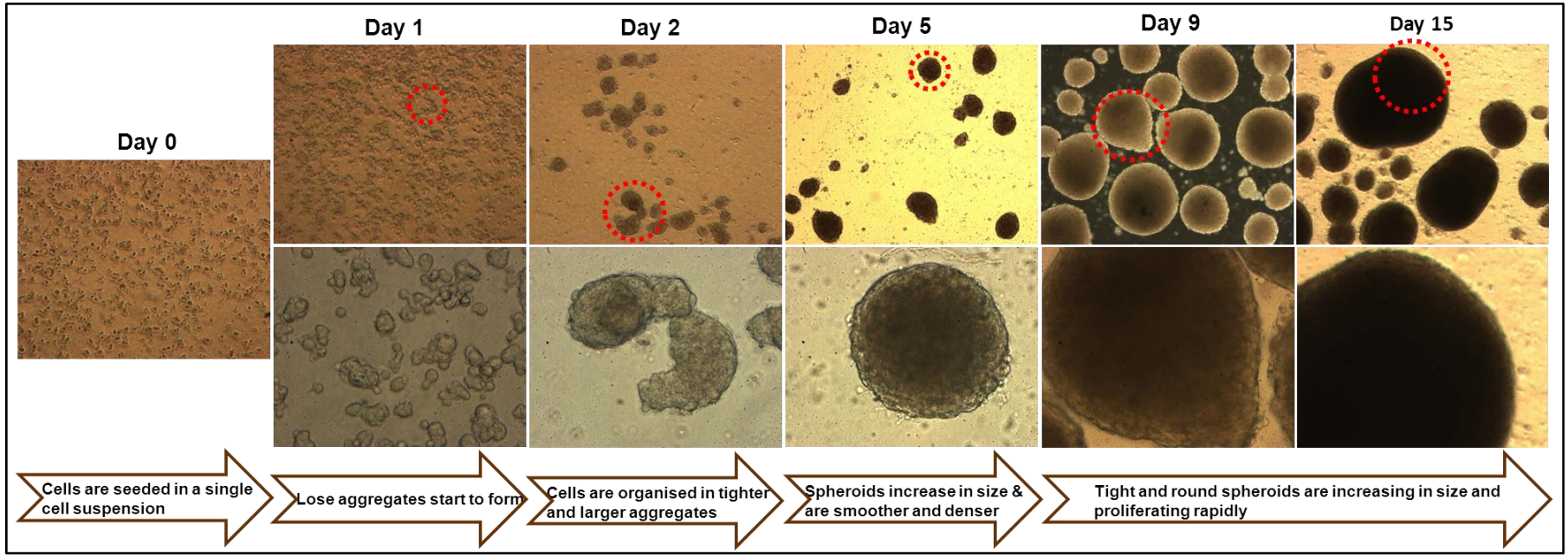


Figure 3.15**:** Formation of FTE 3D cultures in polyHEMA coated plates in 1-15 day time points. Day 0 and top row images were captured under the 50× magnification and bottom row images were captured under the 200× magnification of a light microscope. Visualisation of how the spheroids begin to form as early as in 24 hours after seeding by forming loose aggregates to later proliferate to form larger and more compact spheroids.

.

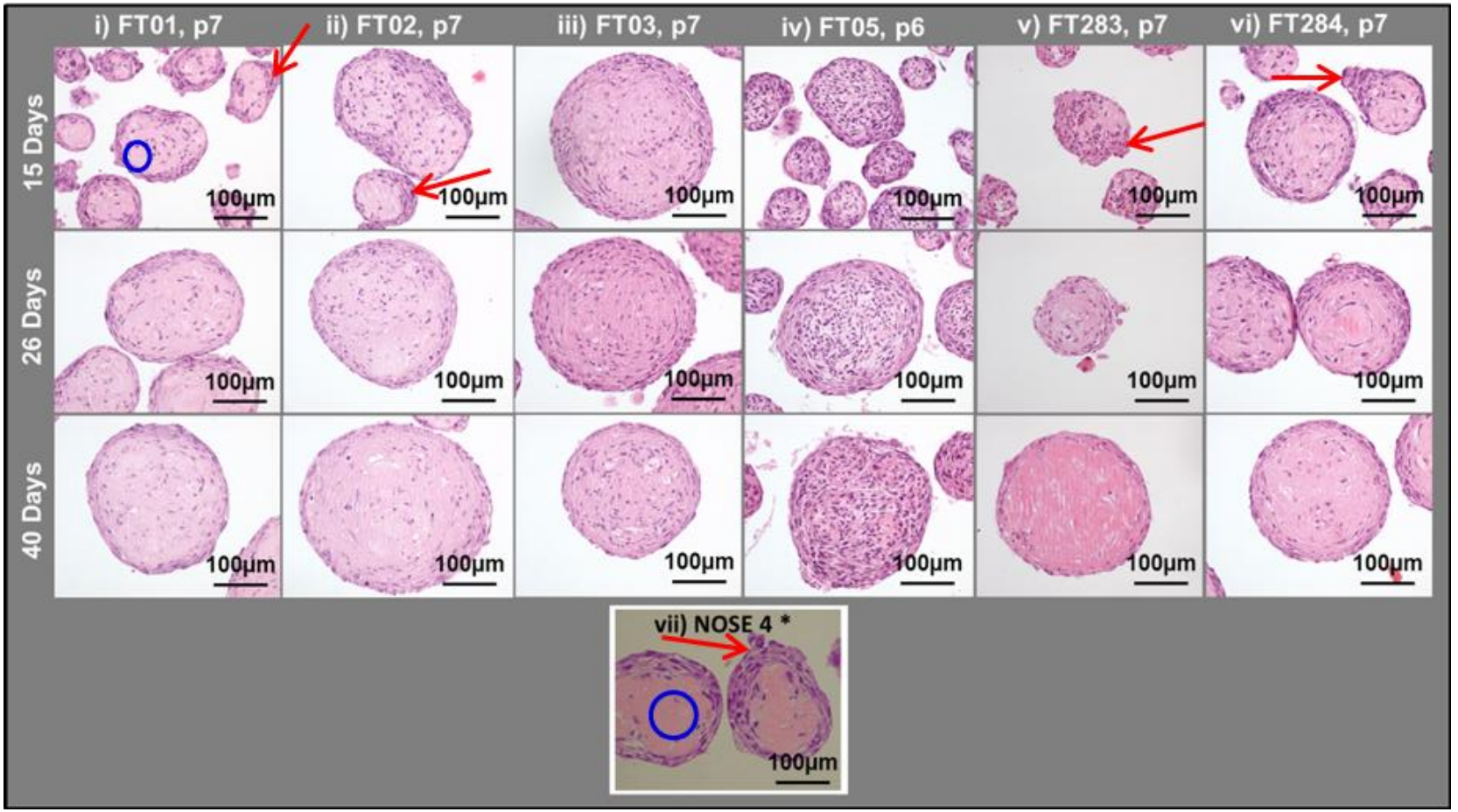


Figure 3.16**:** Representative H&E staining of FTE 3D cultures. 3D cultures of FTE cell lines for 15, 16 and 40 days. H&E stained cross sections of FTE 3D cultures show morphological characteristics and comparison with NOSE 3D cultures. FTE spheroids also present crescentic caps similarly to the NOSE (red arrows), however less prominent that the NOSE. The presence of hyline material is seen in both NOSE and FTE 3D structures (blue circles) but FT spheroids appear to have more cellularity within the cores. (*): NOSE 4 image was contributed by Dr Kate Lawrenson.

3.5 Comparison of FT tissue with 3D FTE and 2D FTE modelling systems.

With the architecture of the 3D FTE modelling system more closely resembling the *in vivo* FT architecture and cell interaction, it was important to compare the expression of key markers between the FT tissue *in vivo*, the 3D FTE cultures and the 2D FTE cultures. FT tissue was collected from the ampullary region of the tube and then embedded into paraffin. 2D cultures cultured on glass chamber slides were also fixed in the same process and stained for several markers to compare with the sections of 3D cultures stained for the same markers.

Four cell lines were taken forward to be stained for several markers: FT02, FT03, FT283, FT284. FT01 was not used to limit the cost of the experiments and because it was the only of the FTE cell lines that had stronger expression of CK7 than AE1:AE3 in contrast with the rest of the cell lines taken forward. FT05 was excluded as it was found to not express any epithelial markers, had abnormal karyotype and aggressive proliferation observed both in the 2D growth curves as well as the differential architecture of the 3D spheroids in exhibited. A summary of the staining for all the selected markers is presented in Table 3.6.

	2D monolayer				3D spheroid				*In vivo*
MARKER	**FT02**	**FT03**	**FT283**	**FT284**	**FT02**	**FT03**	**FT283**	**FT284**	**FT03**
AE1:AE3 (MCK)	++	+++	+++	++++	+/-	+	++++	++	++++
CA125	+	+	+/-	++	-	+/-	++	-	+++
PAX8	+	+	++	+	+/-	+/-	+++	+/-	+++
E-CAD	no data	no data	no data	no data	-	+/-	+	-	+
VIMENTIN	+++	++++	++++	++++	++	+++	+++	+++	++++
LAMININ	++++	++++	++++	++++	++++	++++	++++	++++	++
ER	-	-	++	+	-	-	+	-	+++
MUC-1	++	+	+	++	+/-	-	+++	-	++++
MIB-1 (%)	70%	70%	70%	50%	<5%	<5%	<5%	<5%	<5%
p53Do7	+++	+++	++	+++	+/-	+	+	+/-	+

Table 3.6**:** Summary of immunohistochemical staining comparison between 2D, 3D FTE cell lines and FT tissue. Intensity scoring was performed as previously described for the NOSE immunofluorescence cytochemistry staining. The best model has been the 3D FT283 cell line as it has strong expression of lineage markers that resemble the *in vivo* tissue compared to the 2D culture of this cell line.

3.5.1 Proliferation of FT *in vivo* compared to the 3D and 2D FTE modelling systems

I stained FT tissue and 2D and 3D FTE models with Ki-67 which is an antibody that recognises a nuclear protein coded by *MIB1* gene that is associated with cellular proliferation. It is known that tubal mucosal epithelial proliferation is not evident is normal fallopian tubes. Consistent with this, I found that epithelial cells of the FT mucosa do not have proliferating ability (Figure 3.17). One of the main characteristics of FTE 2D cultures was that they underwent proliferation rapidly in culture which is not a true resemblance of the proliferative abilities of the epithelial cells in the FT *in vivo*. I found that 3D models have significantly reduced expression of Ki-67 which more closely resembles the FT activity *in vivo*.

3.5.2 Epithelial and lineage markers expressed in FT *in vivo* compared to the 3D and 2D FT modelling systems

The epithelial mucosa of the FT *in vivo* expressed the pancytokeratin marker AE1:AE3, and its expression remained strong in the 2D FTE cell lines established. The expression of this marker within the 3D FT spheroids was seemingly lower but was mainly positioned at the outer epithelial layers as seen on Figure 3.18, with the exception of cell line FT283 where a more widespread expression of the marker was observed in the core of the spheroid too. Thus, the outer layer epithelial staining of the 3D cultures resembled more closely the architecture of the FT tissue.

CA125 is developmentally expressed in the Mullerian epithelia including the FT epithelial mucosa (Sharl et al, 1989). However, only weak staining for CA125 was observed in the 2D and 3D FTE cultures (Figure 3.19).

PAX8 is a marker that is expressed by the secretory cells of the FT mucosa and I compared the expression of PAX8 in the 2D and 3D FTE cultures (Figure 3.20). The staining of PAX8 in FT02, FT03, FT283 and FT284 2D cultures was located in the nucleus as expected in 100% of the cells (Figure 3.21). Low levels of PAX-8 staining were also observed in 100% of the cells in the 3D cultures of FT02, FT03 and FT284. 3D culture of FT283 demonstrated a stronger than the 2D nuclear PAX8 staining in ~80% of the cells (Figure 3.21).

E-Cadherin is another Mullerian development key marker expressed in the derived epithelia including the FT epithelium. It is not expressed in OSE due to their uncommitted phenotype but is expressed in the OSE lining of the epithelial inclusion cysts. Immunohistochemical staining for the FT 3D and 2D cultures for this marker was performed. Previously, I showed by immunofluorescent staining that the 2D FTE cultures did not express this marker (Figures 3.7-3.9). E-Cadherin was found very weakly to negatively (+/-) expressed in a small amount of scattered cells within the spheroids, with the FT283 3D culture exhibiting the stronger staining, but still not resembling the strong expression of the tissue *in vivo* (Figure 3.22). This observation might be connected with EMT observed in the morphology of the FT when cultured *in vitro*. Regaining some of the E-Cadherin expression in the 3D cultures demonstrated an ability of the FTE cells to resemble a more epithelial morphology through MET and could possibly indicate a role in the secondary tumour formation from the TIC where MET is important as shown in other cancers as well.

Muc-1 is a glycosylated glycoprotein that is expressed in a variety of epithelial cells *in vivo* including endometrial epithelium and the fallopian tube mucosa. It is up-regulated in the secretory phase of the menstrual cycle and positive staining for Muc-1 is indicative of the perseverance of the secretory cell phenotype and function in culture. Expression of Muc-1 was moderately observed in the 2D FTE cultures and increased Muc-1 expression was only found in the 3D culture of the FT283 cell line (Figure 3.23).

3.5.3 Extracellular matrix markers expressed in FT *in vivo* compared to the 3D and 2D FT modelling systems

Previous work within our group with NOSE 3D cell lines has shown that the expression of extracellular matrix proteins was different between 2D and 3D cultures as cells cultured in 3D tend to form different cell-cell interactions and produce more extracellular matrix proteins (Lawrenson et al, 2009). To test whether the same applied between FTE 2D and 3D cultures I investigated expression for Vimentin and laminin. In contrast to the NOSE 2D vs 3D model, FTE cell lines produce extracellular matrix proteins Vimentin and laminin when cultured in 2D and 3D (Figure 3.24 and 3.25)

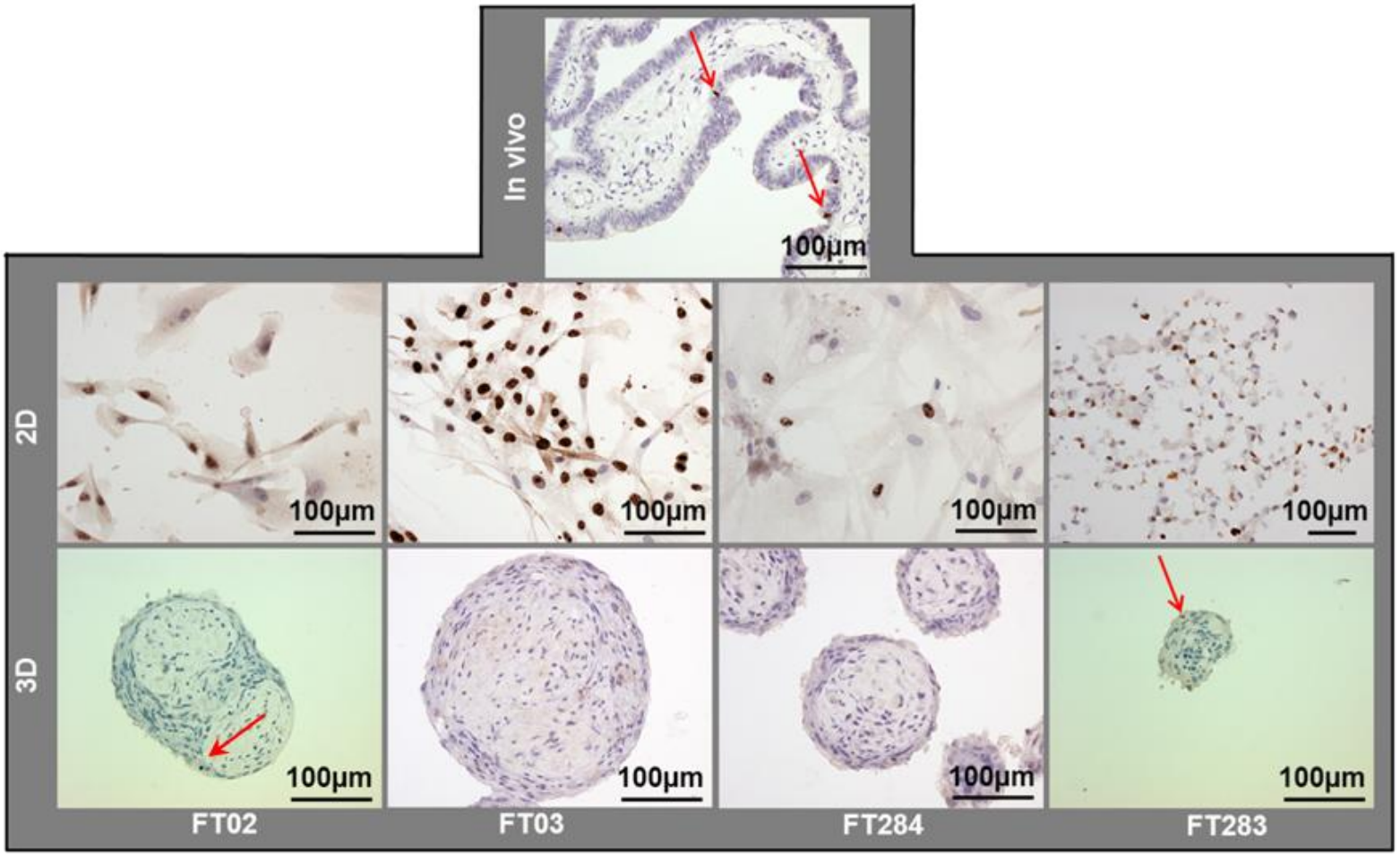


Figure 3.17**:** Immunohistochemical staining for proliferation of FT *in vivo* compared to 2D and 3D FTE cultures. Samples were stained with Ki-67. The *in vivo* section is showing tubal fimbria where papillary processes are covered by tubal epithelium with a core of fibrovascular tissue and vessels. At the tubal epithelium normal architecture is observed with the elongated ciliated cells with the more rounded interspersed secretory cells. The tubal epithelium is multi-layered with cellular stratification similarly to the multilayered 3D spheroids of the FTE cultures. The arrows show the some of the cells that express Ki-67 in the tubal mucosa and the 3D cultures. It is evidence that the expression of Ki-67 is abundant in the 2D FTE cultures, especially for cell lines FT02, FT03 and FT283 and significantly reduced in the 3D cultures.

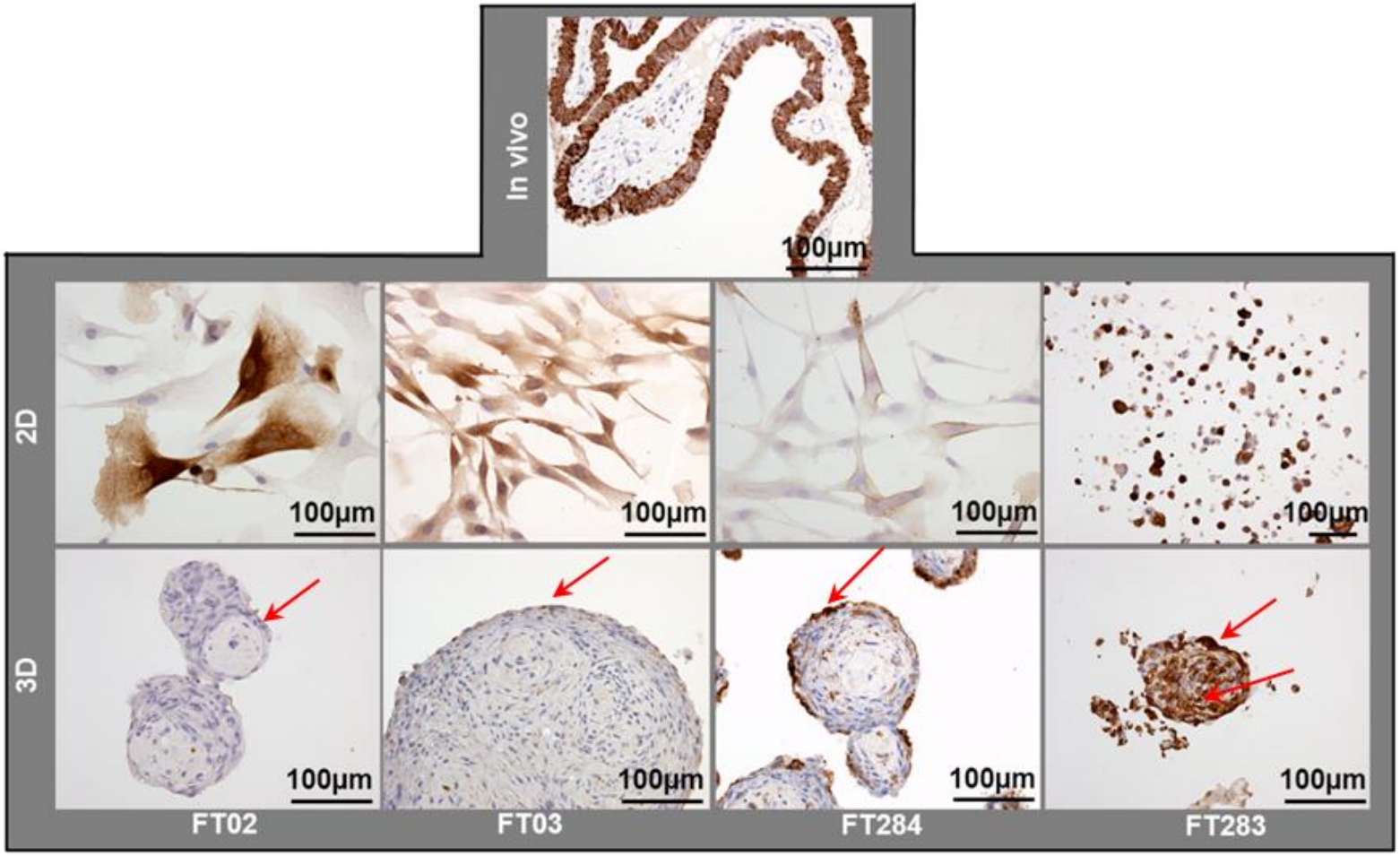


Figure 3.18: Immunohistochemical staining for AE1:AE3 epithelial marker compared between FTE cell lines in 2D, 3D and FT tissue. Universal staining observed in the cytoplasm of all the 2D cultures and outer layer staining for 3D structured for FT02, FT03 and FT284.

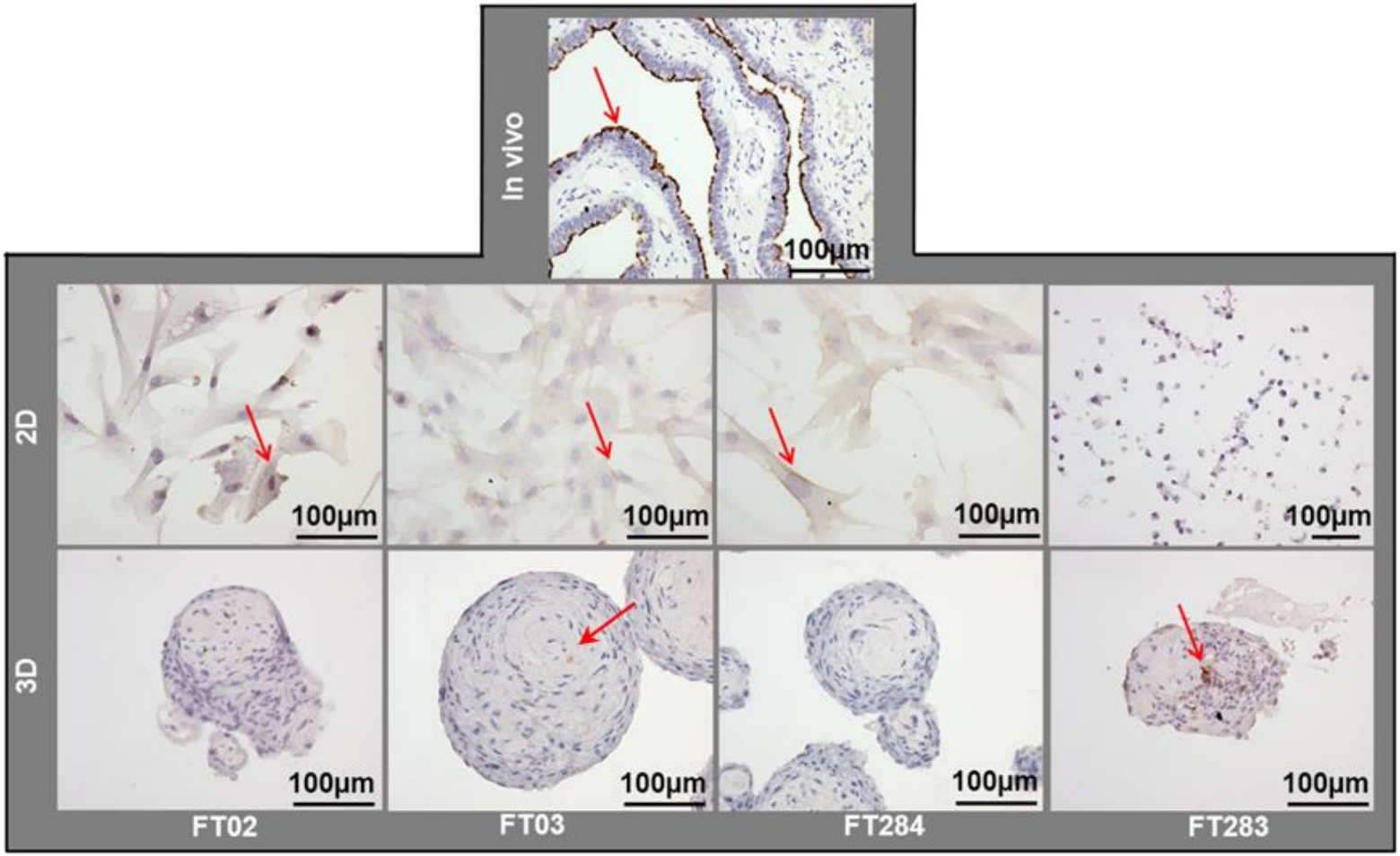


Figure 3.19: Immunohistochemical staining for CA125 lineage marker compared between FTE cell lines in 2D, 3D and FT tissue. The red arrows show representative areas of positive cytoplasmic staining. Increased expression of CA125 was found for FT283 when cultured in 3D.

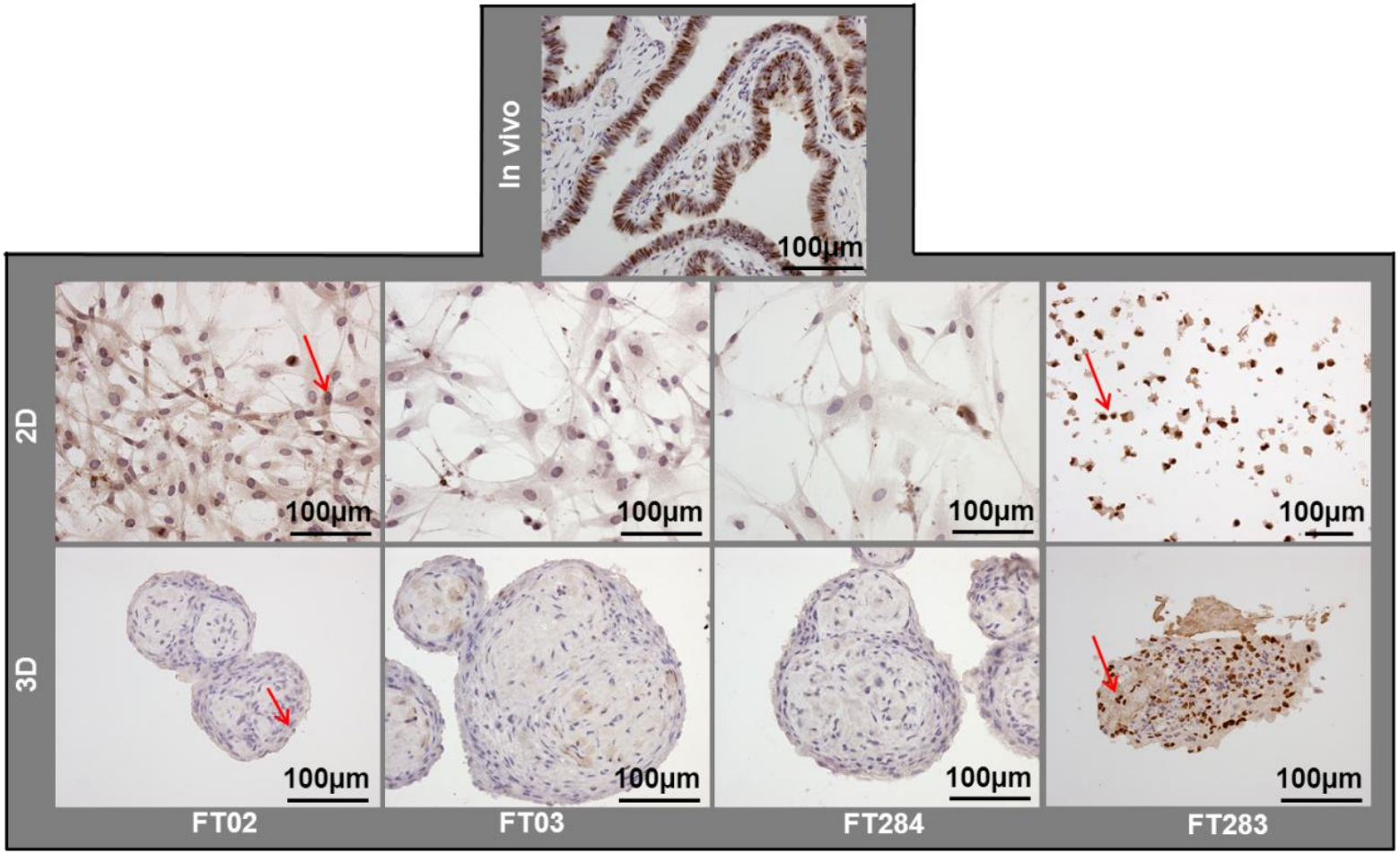


Figure 3.20**:** Immunohistochemical staining for PAX8 lineage marker compared between FTE cell lines in 2D, 3D and FT tissue. Striking difference between the staining of FT283 3D and the rest. FT283 seems to be a complex structure comprising of a mixture of secretory cells strongly expressing PAX8 and presumably ciliated cells that are not expressing PAX-8. The brown cytoplasmic staining is unspecific staining due to staining conditions.

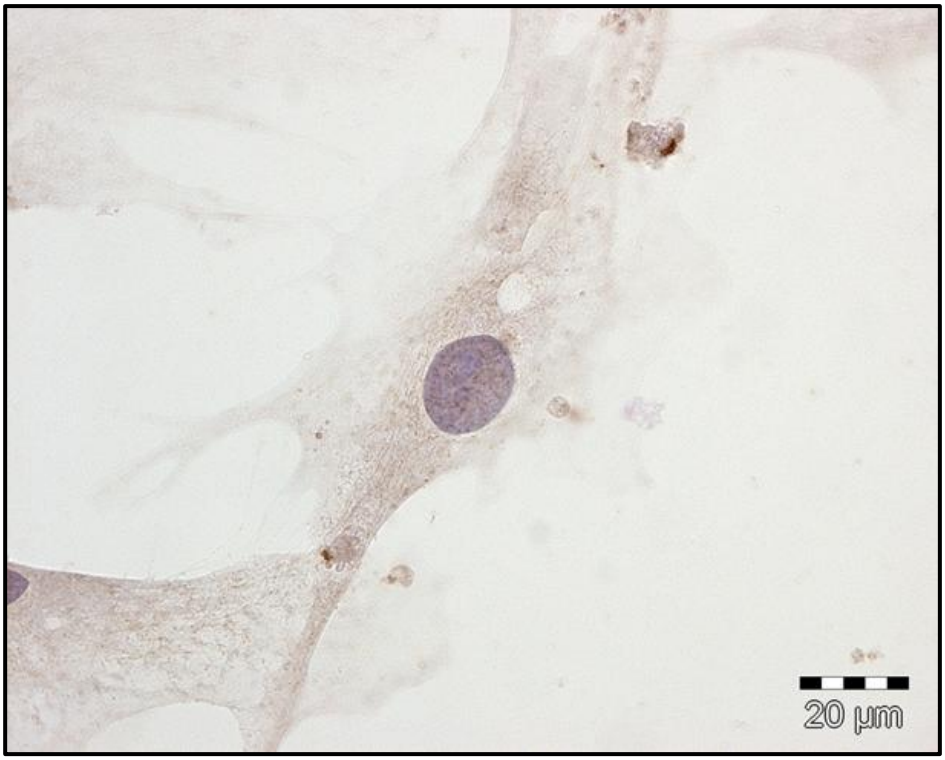


Figure 3.21**:** Immunohistochemical staining showing a punctate nuclear PAX8 staining in a 2D FTE culture. In the majority of the FT 2D cultures a faint nuclear staining for PAX8 was observed in 100% of the cells in culture.

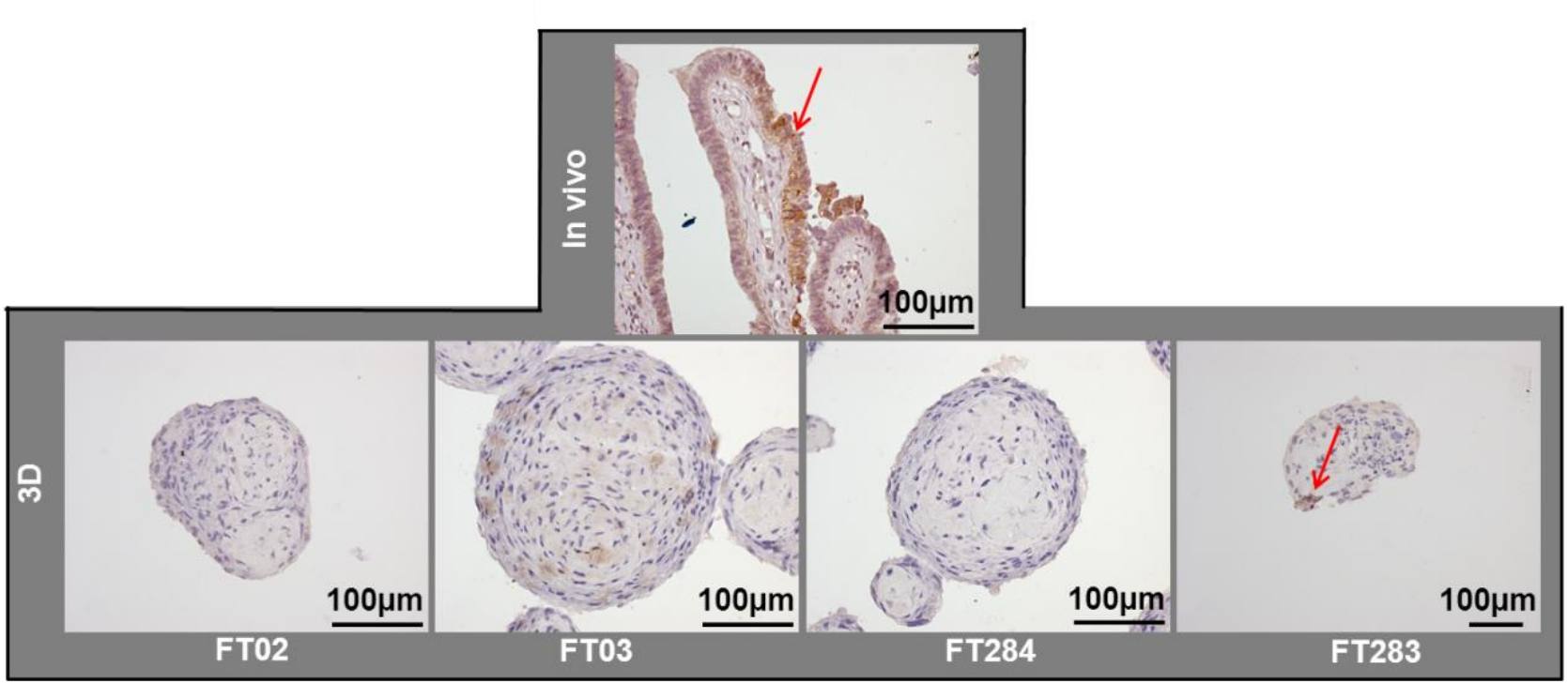


Figure 3.22**:** Immunohistochemical staining for E-Cadherin marker compared between FTE cell lines in 3D and FT tissue. The arrows show examples of positive staining.

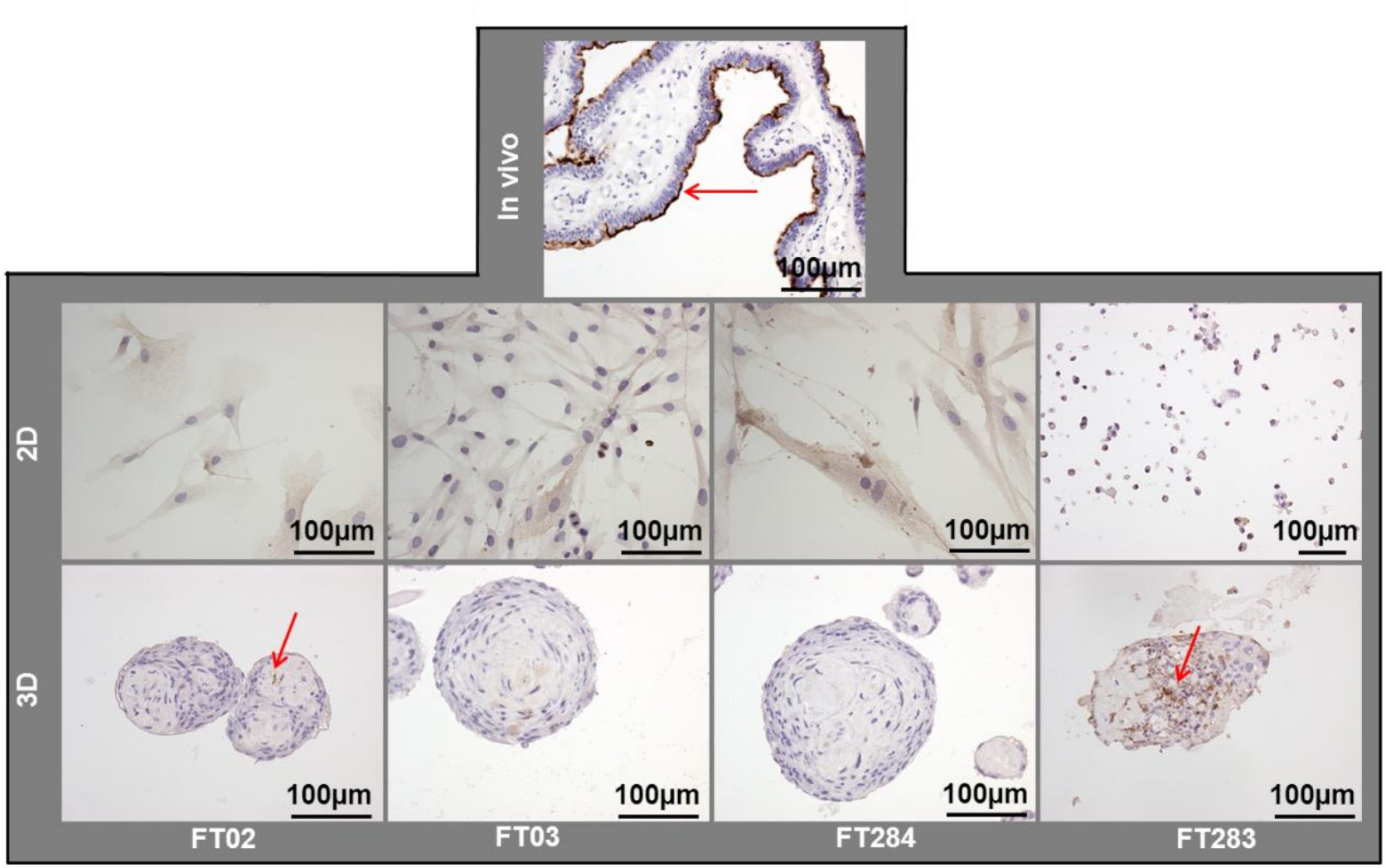


Figure 3.23**:** Immunohistochemical staining for MUC-1 marker compared between FTE cell lines in 2D, 3D and FT tissue. Increased expression of MUC-1 was found for FT283 when cultured in 3D.

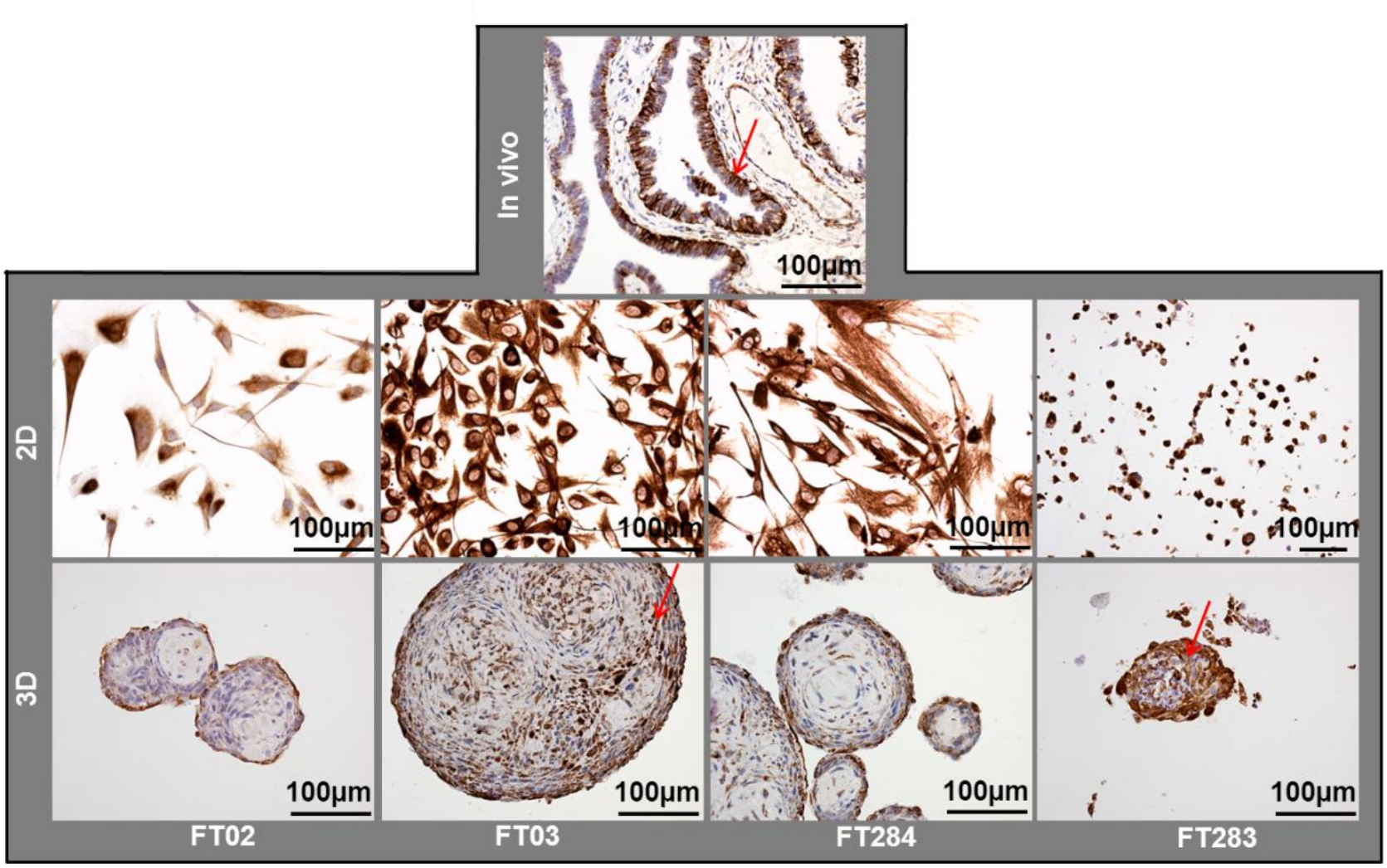


Figure 3.24**:** Immunohistochemical staining for Vimentin ECM marker compared between FTE cell lines in 2D, 3D and FT tissue. The red arrows show the localisation of Vimentin expression in the basement membranes between the FT mucosa cells in vivo and a similar localisation observed in the matrix produced by the FTE cells in the 3D cultures.

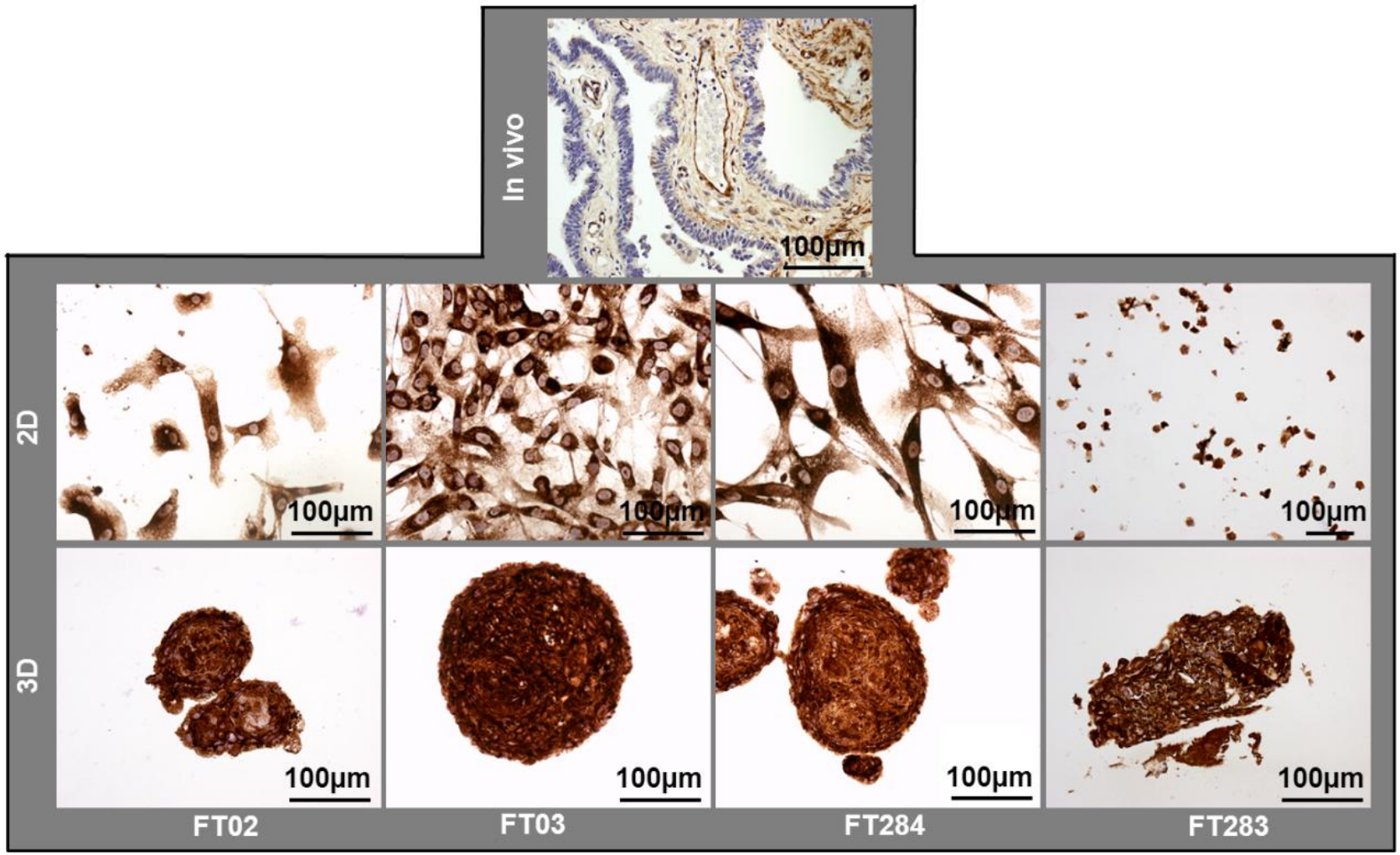


Figure 3.25**:** Immunohistochemical staining for laminin marker compared between FTE cell lines in 2D and 3D and FT tissue. Laminin is expressed by FTE cells equally in 2D and 3D.

3.6 Evidence supporting a dual cell of origin for EOC development

3.6.1 PAX8 evidence supporting a dual origin of EOC

As previously mentioned it is proposed that EOC may have a dual origin from FTE and NOSE cells. To investigate the dual cell origin hypothesis for EOC I have investigated the differential expression of *PAX8* between NOSE, FTE and EOC cell lines and evaluate the expression for this lineage marker. Briefly, NOSE and FTE cell lines that have been confirmed to be normal after histopathology examination and that have been confirmed to have an epithelial status (n=60 and n=5 respectively) were cultured and RNA was extracted from cell lines of passage 3 to passage 8. RNA was also extracted from 47 commercially available EOC cell lines shown in Appendix 1, Table 3. The extracted RNA's quality was confirmed with spectrophotometry and agarose gel electrophoresis and subsequently transcribed into cDNA as described in the methods section. The differential expression was assayed performing a Fluidigm gene expression assay. The study design and the quality control analysis will be described in more detail in Chapter 5.

Linear regression analysis was performed to observe if there was a significant difference in the expression of PAX8 between the 3 types of cells. I found that *PAX8* was over-expressed in EOC cell lines compared to NOSE and FT and that there was a trend of increasing *PAX8* expression from NOSE to FT to EOC cell lines ($P=4.5\times10^{-6}$, Figure 3.26 A). This result should be taken with caution as only one FTE cell line has been in included in the analysis, the rest of the samples failed to amplify. It is worth mentioning that this cell line is the FT283 which has been shown to have the highest levels of PAX8 in the immunohistochemistry staining experiments.

I have also investigated the expression of *PAX8* between Serous Ovarian Carcinomas (SOC) and normal FT using mRNA expression data publicly available from the TCGA (www.cbioportal.org.uk) and found consistent with my observations that PAX8 was significantly over-expressed in SOC compared to normal FT (P=0.002, Figure 3.26 B). *PAX8* has also been found amplified in high grade EOCs (TCGA, 2010).

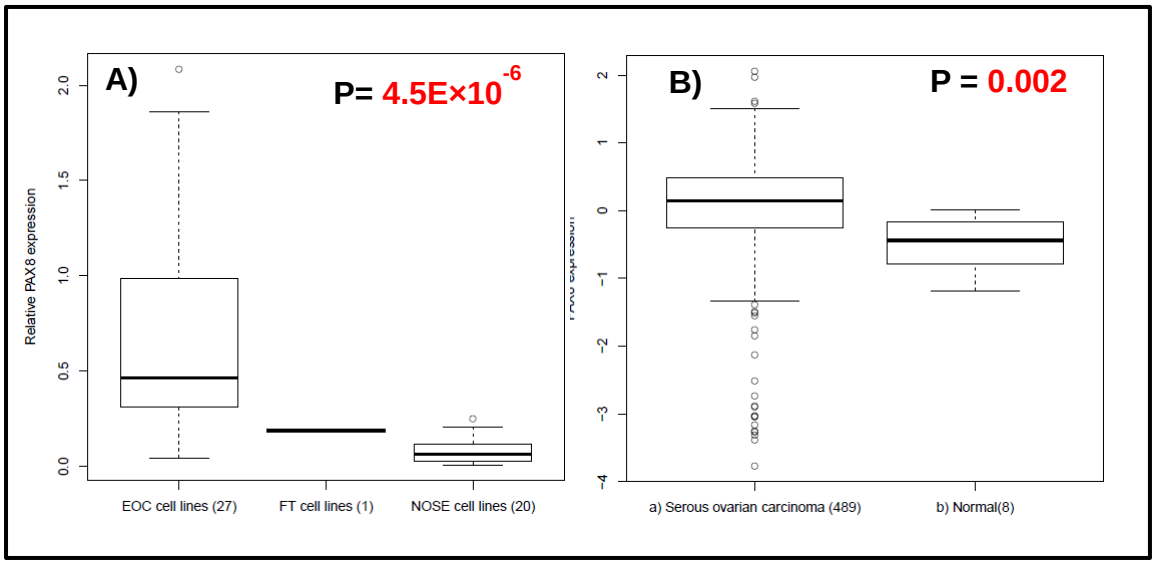


Figure 3.26: Boxplots representing PAX8 differential expression. A) Differential expression of PAX8 between EOC, FT and NOSE cell lines. B) Differential expression of PAX8 between SOC and Normal FT according to TCGA

To date PAX8 has been shown to be expressed in thyroid follicular cells and other cells of the reproductive system. It was intriguing to investigate whether this reproductive lineage marker was expressed in NOSE cell lines and if so to what extent. I stained cell lines NOSE 231 L, NOSE 228 R and NOSE 270 L and found that the expression of PAX8 varied in those cell lines with one of them expressing strong PAX8 in about 20% of the cell population and the other 2 expressing low PAX8 in 10% of the cell population (Figure 3.27).

I found that 50% of the patients (n=6) exhibited expression of PAX8 in the OSE *in vivo*. My study also confirmed previously reported expression of PAX8 in inclusion cysts and identified expression of PAX8 in other ovarian structures such as in rete ovarii and mesoneprhic tubules and ducts (Table 3.7, Figure 3.28).

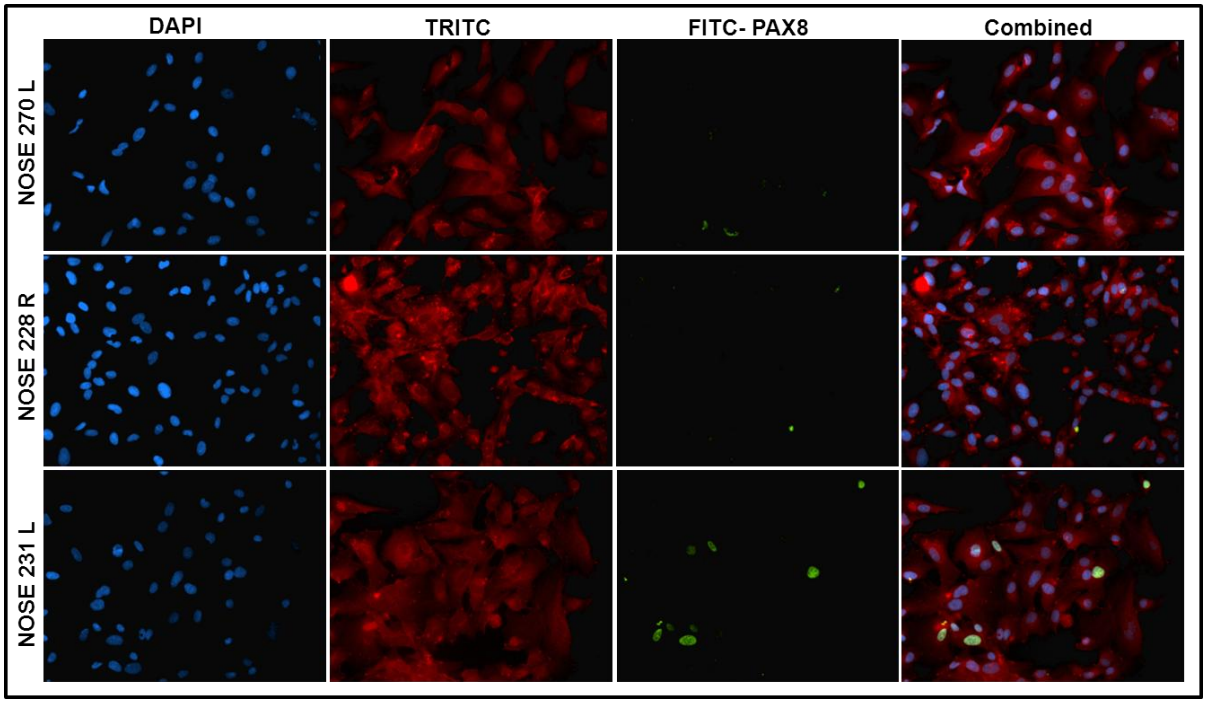


Figure 3.27**:** Fluorescent immunocytochemistry of NOSE cell lines for FT and reproductive tract lineage marker PAX-8 (Images at 200×). The blue stain is DAPI enabling the visualisation of the nuclei. The red stain is staining the cytoplasm. Finally, the green stain is indicating expression of PAX8 which is expressed by 10-20% of the cells in culture, showing strong expression in only 1 of the 3 cell lines. The positive and negative controls for PAX8 are shown in figure 3.10.

Patient number	**Age at surgery**	**OSE staining**	**Other PAX8 +ve structures**
N1	56	-	Not present
N2	65	++	Inclusion cysts Rete ovarii
N3	67	+++	Inclusion cysts Rete ovarii Mesonephric tubules
N4	75	Not present	Paraovarian cyst Inclusion cyst
N5	62	++	Mesonephric duct
N6	74	-	Inclusion cysts
N7	57	-	Not present

Table 3.7**:** Summary of immunohistochemical staining for PAX8 expression in the OSE and cortical inclusion cysts in normal ovaries.

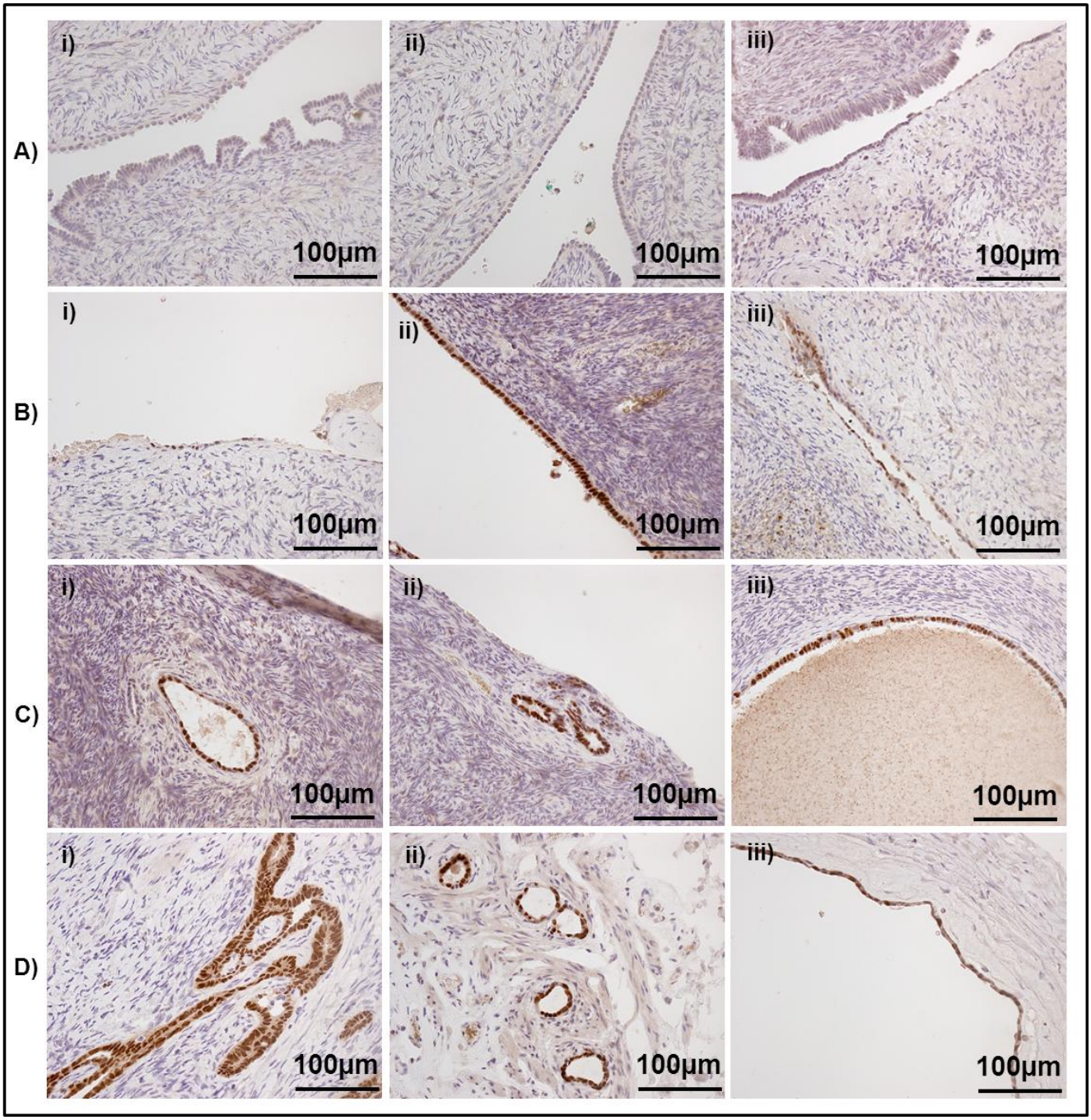


Figure 3.28**:** Immunohistochemical staining showing expression of PAX8 in inclusion cysts and ovarian surface epithelium. A) OSE staining negative for PAX8: i) N7 OSE, ii) N1 OSE, iii) N6 OSE. B) OSE staining positive for PAX8: i) N2 OSE, ii) N3 OSE, iii) N5 OSE. C) Ovarian inclusion cysts staining positive for PAX8: i) N2 inclusion cyst, ii) N3 early inclusion cyst, iii) N6 inclusion cyst. D) Other ovarian structures staining positive for PAX8: i) N2 rete ovarii, ii) N3 mesonephric tubules, iii) N4 paraovarian cyst.

3.6.2 Assaying differential expression of genes implicated with EOC development in NOSE, FTE and EOC cell lines.

An important aim of this thesis was to create a large repository of NOSE cells and assay their differential expression to EOC cell lines in order to identify candidate genes involved with EOC development emerging from large screening studies. Given the dual origin hypothesis for EOC development a small panel of FTE cell lines will be included to form together with the NOSE cell lines the normal cell lines that the expression of candidate genes will be compared to the expression in EOC cell lines in the following chapters.

In order to validate the NOSE& FTE versus EOC differential expression as an appropriate model to identify or validate genes emerging form large screening studies, the expression of genes that have a well-established in EOC development was investigated. Six genes were investigated due to their prominent role in propagating EOC development using Taqman gene expression assays in the Fluidigm 96.96 Dynamic array. The design of this assay will be discussed in more detail in chapter 5 of this thesis. The mRNA expression data from NOSE & FTE versus EOC cell lines were compared with publicly available microarray data from the TCGA. The TCGA database now includes publicly available expression data for several of the common human cancers. mRNA expression data from 489 serous ovarian carcinomas (SOC) and 8 normal fallopian tube (FT) tissues were downloaded from CBIO portal (http://www.cbioportal.org). It should be noted that the normal FT samples in TCGA are samples coming from whole normal fallopian tubes. Thus, some stromal contamination could be present. The analysis was performed in 2 groups of datasets. The first dataset contained mRNA expression data for all genes coming from median values from all three mRNA expression platforms used. The second dataset contained merged and scaled mRNA expression values of genes represented by all three expression platforms. In the present study I will present only the analysis using the first dataset as it was the more comprehensive and provided data for the highest number of samples. I have assessed any possible differences in the expression of the candidate genes in SOC and normal samples from TCGA using the non-parametric Wilcoxon rank sum test in R software.

Three genes that are well established oncogenes in EOC development are *BRAF*, *KRAS* and *MYC*. Linear regression analysis was performed to assay the differential expression in NOSE, FTE and EOC cell lines. I found that *BRAF* and *KRAS* were over-expressed in EOC cell lines compared to NOSE and FTE cell lines and that there was a trend of increasing expression for both genes from FT to NOSE to EOC cell *lines* ($P= 4.22\times10^{-5}$ and $P= 0.019$ respectively, Figure 3.29 A & C). *KRAS* and BRAF were also found over-expressed in SOC compared to normal FT tissue according to TCGA ($P= 2.09\times10^{-5}$ and $P= 2.53\times10^{-4}$, Figure 3.29 B &D). A small trend of increasing *MYC* expression was also observed. *MYC* expression was found to increase from NOSE to FT to EOC cell lines ($P= 0.021$). The expression of the gene was not differential between SOC and normal FT according to TCGA (Figure 3.29 E & F).

Three tumour suppressor genes implicated in EOC development have also been investigated, *BRCA1*, *BRCA2* and *TP53*. The results were not as expected, interestingly not confirming the loss of function role these genes have in EOC development. All three genes were found to be significantly over-expressed in EOC compared to both NOSE and FTE cell lines (Figure 3.30 A, C & E). Consistent with this observation a trend of over-expression for *BRCA1* was found in SOC compared to normal FT according to the TCGA mRNA data, although not statistically significant. The same trend of over-expression in SOC compared to normal FT was found for *BRCA2* and that was statistically significant ($P= 1.38\times10^{-6}$, Figure 3.30 D). The observed and unexpected over-expression of EOC implicated tumour suppressor genes in EOC compared to NOSE & FT also observed for some of them in the SOC versus normal FT model will be further discussed.

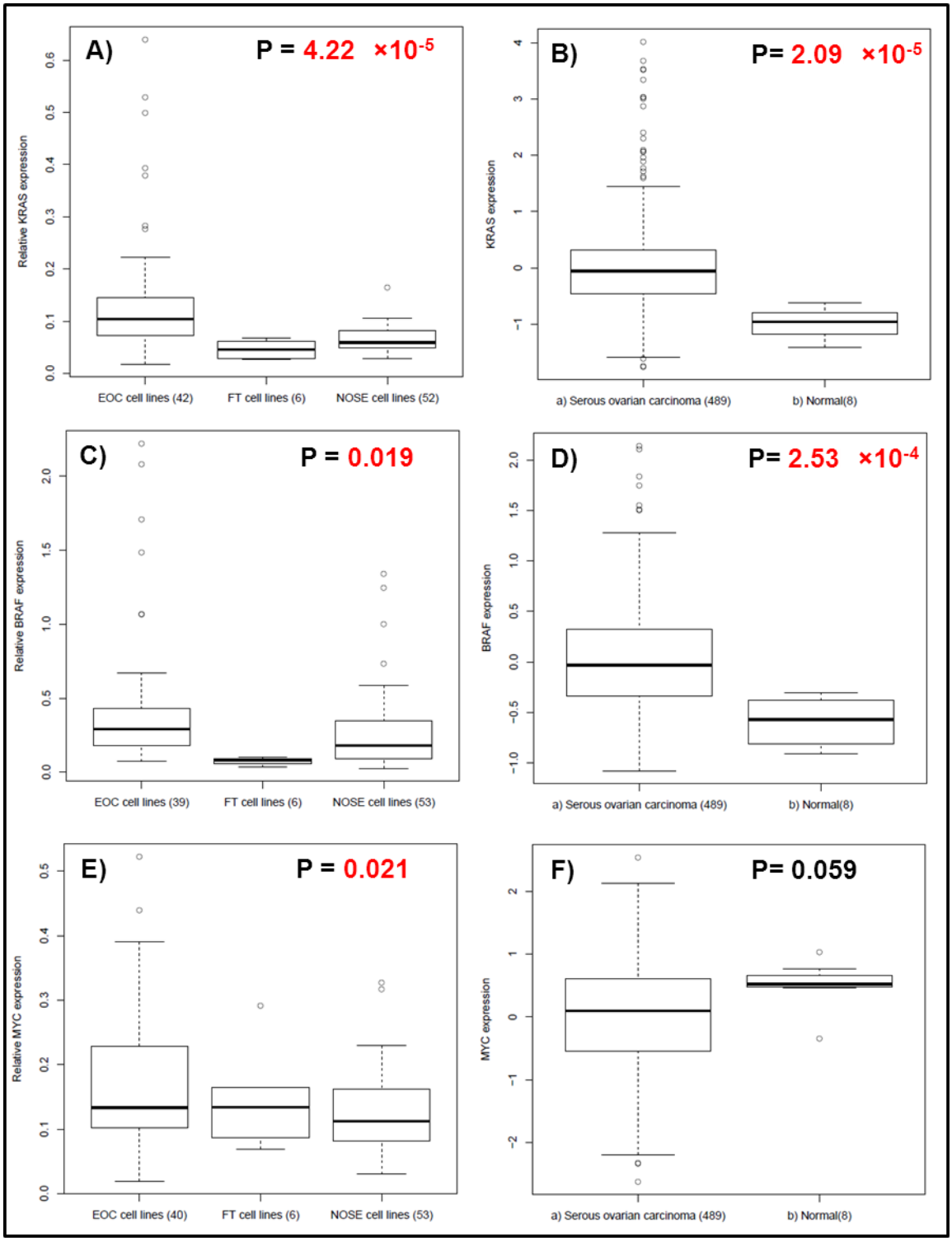


Figure 3.29**:** Differential expression of A) *KRAS, C) BRAF* and E) *MYC* between NOSE, FTE and EOC cell lines. Increased expression of these genes was observed in EOC cell lines, which is in line with the role that they have in EOC development. B) KRAS and D) *BRAF* over-expression was also observed in SOC compared to normal FT but no differential expression was found for F) MYC in SOC and normal FT according to TCGA mRNA expression data analysed.

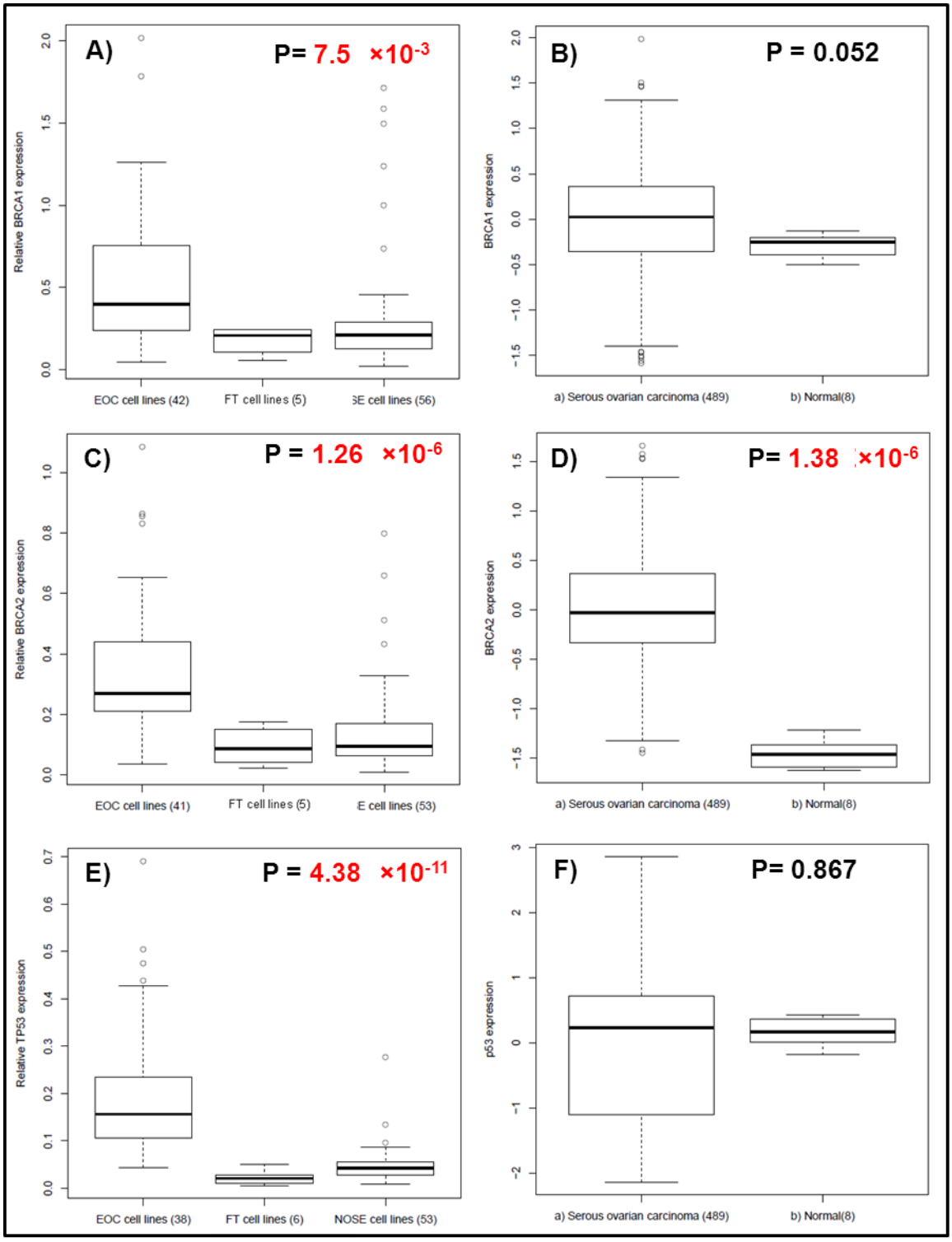


Figure 3.30**:** Differential expression of A) *BRCA1*, C) *BRCA2* and E) *TP53* between NOSE, FTE and EOC cell lines. Increased expression of these genes was observed in EOC cell lines. In line with this observation, B) a trend for over-expression of *BRCA1* in SOC compared to normal FT was observed and D) *BRCA2* was significantly over-expressed in SOC compared to the normal samples according to TCGA.

3.7 Discussion

3.7.1 Establishing a NOSE cell line repository

Understanding the biology of EOC development is closely interlinked with improving the design of experimental approaches used to evaluate and validate biomarkers associated with EOC development. Several studies have assayed the differential expression of genes between EOC and normal ovaries in order to identify genes associated with early stage disease and development. Several of those studies have used microarray analysis using RNA from EOC tumour samples and bulk normal ovaries. Over the years several genes have been reported to be differentially expressed in serous ovarian carcinomas and normal ovaries using microarray analyses (Ono et al, 2000, Welsh et al, 2001, Grisaru et al 2007). These studies used RNA from EOC tissue and whole normal ovaries to compare mRNA levels and then proposed to have identified several candidate molecular markers for EOC development. Whole normal ovaries, however, have a mixture of epithelial, stromal, endothelial and other types of cells. Therefore it is possible that the differential expression found between bulk extracted normal ovary and ovarian cancer samples might partially reflect the expression differences between the cell populations rather than a real implication of the genes in neoplastic development. Therefore, it would be more appropriate to compare the mRNA levels in EOC samples with the mRNA levels of the proposed precursor cells of epithelial origin NOSE and FTE cells. In realisation of the need to evaluate differential expression between EOCs and the specialised precursor cells rather than whole ovaries several small scale studies have used small numbers of NOSE or IOSE cell line RNA or ovarian surface epithelial scrapings instead of normal ovaries to assay differential gene expression (Mok et al, 2001, Tonin et al, 2001, Zorn et al, 2003, Huddleston et al, 2005).

Although NOSE cells have been previously isolated, cultured and characterised (Auersperg et al, 1999, Lawrenson et al, 2009), the value of this study lies with it being the first study that has established a large repository of NOSE cell lines in order to be used for assaying the differential expression of candidate genes with a potential role in EOC development compared to EOC

cell lines. All of the collected cell lines expressed the epithelial markers CK7 and/or AE1/AE3 and only one was excluded due to a possible stromal contamination. Some of the established primary NOSE cell lines showed small levels of FSP expression but was much lower than the expression of the marker by the control stromal cell lie INOF2 used. Additionally, It is well established that OSE cells have an uncommitted phenotype *in vivo*, expressing both epithelial and mesenchymal markers (Auersperg et al, 2001) and that they show this phenotypic plasticity *in vitro*, where they undergo epithelial to mesenchymal transition and take up a more fibroblastic like morphology (Salamanca et al, 2004, Zavadil and Böttinger, 2005, Ahmed et al, 2006).

Alongside the characterisation of NOSE cell lines for epithelial status I made an observation that the expression of pan-cytokeratin was decreased as the age of the patients increased but was not statistically significant. However, the trend observed could further be investigated in a real time expression analysis where expression of specific cytokeratins correlated with age of the individuals could be assayed. To my knowledge there are no reports that have shown an association between cytokeratin expression and the age of the patient in any epithelial cell type. One previous study has investigated in the expression of cytokeratins relative to the age of the patients in seminoma cases of the human testis and found no association (Cheville et al, 2000).

3.7.2 Establishing primary FTE cell lines in 2D and 3D and comparing to FT tissue

Given the recent proposal for multiple cells of origin for EOC, it was hypothesised that it is necessary to study EOC tumorigenesis using both NOSE and FTE *in vitro* models. Thus, I aimed to establish and characterise primary FTE cell lines and also create a 3D model of these cultures hypothesising that it would better reflect the architecture and biological activity of FTE cells in vivo. Successful isolation and sub-culturing of 6 FTE cell lines was achieved. One of them was found to have an abnormal karyotype possibly indicating that it may be a transformed cell line and that could explain the differences observed in its morphology and staining patterns compared to the rest. This cell line was FT05,

and its abnormal karyotype could be related to the aggressive proliferative phenotype this cell line showed and the absence of epithelial marker expression.

An interesting finding was the observation that FTE cell lines formed colonies in soft agar even when seeded in low densities (5000 cells/well). This observation was initially surprising as ability for anchorage independent growth is mainly attributed to transformed cell lines. One speculation would be that the FTE colonies on soft agar were aggregates of cells formed due to their ability to migrate. However, previous work has demonstrated that the ability of FTE cells in *ex vivo* cultures for healing in wound assays was due to detainment of their proliferative ability rather than migration (Levanon et al, 2010. The colony like structures I observed were not very cellular and produced a hyline type material typically observed in 3D cultures. This may suggest that in soft agar FTE cells form colonies not due to being transformed but more due to an ability to form 3D structures that recapitulate the microenvironment of the tissue accompanied with the production of extracellular matrix. The FTE cells look polarised in some of the larger structures, arranged on the outer layer of the colony like formations. Further work to investigate the polarity of those cells in those structures would be interesting to follow.

The FTE cell lines expressed epithelial markers CK7 and AE1:AE3 with one being more expressed than the other in certain cell lines. It is conceivable that the tubal epithelial mucosa is diverse across the tube regarding the expression of epithelial markers. The procedure collection was kept as constant as possible but depending on the part of the tube that was given by the surgeon in surgery the brushing of the mucosa could be closer to the fimbrial end or the ampullary region.

It has been previously shown that *ex vivo* FTE cultures expressed BeREP4 and E-Cadherin (Levanon et al, 2008) which was not the case with the *in vitro* 2D cell cultures I maintained. It is possible that different collection methods in *ex vivo* and *in vitro* systems can lead to expression and maintenance of different markers.

An interesting observation was the absence of CA125 expression in all the cell lines apart from FT02. CA125 would be expected to be expressed in

FTE cells given the Mullerian developmental origin of FTE cells. Clinically, the only difference between FT02 cell lines and the rest was that it was derived from a patient had some evidence of chronic salpingitis although the cell line was found to be karyotypically normal and was included in subsequent analyses. Previous work has shown that patients with acute chronic salpingitis have elevated blood levels for CA125 (Moore et al, 1998) although this may not necessarily relate to the FT02 cell line specific levels of CA125. All the FTE cell lines were found to stain positive for Vimentin which could be explained as these cells have phenotypic plasticity *in vitro* and their morphology is rapidly changing to more epithelial-mesenchymal as they are sub-cultured further, a process known as EMT.

Previous work performed in an ex-vivo FT system has established that an FT secretory cell lineage marker is PAX8 (Levanon et al, 2010). I found that 90-100% of the cells in all the primary FTE cultures I established expressed PAX8 indicating that they are rich in FT secretory (FTSEC) cells. This is an indication that secretory cells' lineage characteristics are maintained in these *in vitro* cultures with the ciliated cells reverting to a phenotype of the secretory cells or that culturing conditions are not optimal for ciliated cell growth.

FTE cells were further cultured in 3D *in vitro* models. Previous work within our group has established that the architectural organisation and certain molecular features of the OSE *in vivo* are better represented by NOSE 3D cultures compared to the 2D counterparts (Lawrenson et al, 2009). The most striking observation was the restoration of the FT *in vivo* architecture and cell interaction with extracellular matrix in the 3D cultures with the production of the hyline material. I also found that the FTE 3D cultures are better representing the true biological activity of the FT epithelial mucosa *in vivo* mainly due to the universal reduction in the expression of proliferative marker Ki-67 and cell cycle regulator p53 in the 3D cultures compared to the 2D cultures. FT epithelial mucosa proliferation is not normally found in normal FT tissue *in vivo* and if found it is mild. It has been reported that elevated tubal proliferation does not exceed 4% of the 72 normal tubes tested (Yanai-Inbar et al, 2000). Since tubal epithelial proliferation in the literature is associated with the presence of malignant lesions (Yanai-Inbar et al, 2000) an FTE model that recapitulates the

tissues proliferative activity such as the 3D cultures I generated will be more appropriate for future studies of EOC carcinogenesis for subtypes originating from FTE epithelial cells. Additional work within our group has investigated using gene microarrays gene expression differences between the 2D and 3D FTE cultures. Interestingly, it was found that the less proliferative phenotype of 3D FTE cultures was reflected in global gene expression changes that occurred when FTE cells were transferred from a 2D to a 3D microenvironment. Genes associated with cell cycle regulation, DNA replication and DNA repair were found to be down-regulated in the 3D FTE cultures possibly responsible for the observed reduction in proliferation observed (Lawrenson et al, *submitted and under review*).

The expression of PAX8 and CA125 was increased in 3D cultures compared to the 2D FTE cultures and better representing their expression found *in vivo* only for cell line FT283 so that was not a universal observation. It was surprising that the 3D models seemed to generally express more moderately the lineage marker PAX8. Similarly, MUC-1 which is a marker that represents the secretory phase of FTE cells was strongly expressed in the 3D culture of FT283 cell line compared to the 2D but this not a consistent observation for all the cell lines. A possible explanation could be that when the cells were cultured in 3D, a proportion can reverse back to the ciliated phenotype that was lost in the 2D cultures. Assuming that is true, I propose that the FTE 3D models I established would still be an appropriate model to study EOC development because even if some serous EOCs are mainly proposed to originate from FTSEC, interaction of ciliated and secretory cells together with the influence from the tissue microenvironment might be important for neoplastic development.

3D cultures seemed to produce a matrix where laminin and Vimentin are major components but the expression of both markers was not increased notably compared to the 2D cultures. However, the organisation of the extracellular matrix in the 3D cultures resembled more closely the *in vivo* expression and localisation than 2D cultures. Vimentin staining at the 3D spheroids was observed only in between cells, indicating the importance for it to maintain the structure of these spheroids and recapitulating the *in vivo*

organisation. Finally, the organisation of the epithelial marker AE1:AE3 expression in the outer layer of the FTE 3D cultures resembled more closely the *in vivo* localisation that the expression in the 2D cultures.

3.7.3 Investigating EOC using combined *in vitro* models of NOSE and FTE

This study was based on the hypothesis that EOC has a dual if not multiple cell of origin. Previous work has shown that the expression of the Mullerian lineage marker CA125 is present in FTE tissue, OSE lining of ovarian inclusion cysts and EOC tumours but not in the OSE *in vivo* (Auersperg et al, 2001). Thus, I investigated the expression of *PAX8*, another lineage marker, between FTE, NOSE and EOC cell lines using mRNA from the established primary NOSE and FTE cell lines and commercially available EOC cell lines. Expression of *PAX8* was found to be higher in FTE than NOSE and higher in EOC than both NOSE and FT, but that was based only in one FTE sample. Assuming this would be validated, based on that observation, it is conceivable that ectopic relocation of FTE cells to the ovary could be responsible for the development of high grade serous EOC. Alternatively, tumours may develop in the tube and involve the ovary later on since most EOC are diagnosed at a late stage. This is in line with reports that CA125 that is the stronger EOC marker is not expressed in NOSE cell lines but is in FTE. Supporting this model, a recent study has found that PAX8 is expressed in mucinous EOC tumours created from FTE cells that were subjected to ectopic expression of human HRAS (Shan et al, 2012). *PAX8* was also reported to be amplified in high grade serous ovarian carcinomas according to the TCGA (The Australian Ovarian Cancer Study Group/Australian Cancer, Study, 2011).

Interestingly, I found that PAX8 was expressed in low levels in the NOSE cell lines too (10-20%) of the cultured cells. I have also found that PAX8 was expressed in OSE, the OSE lining of ovarian inclusion cysts and other ovarian structures *in vivo.* There are previous reports that have indicated that PAX8 expression is present in the epithelia of the ovarian inclusion cysts but not in the OSE (Bowen et al, 2007). My observation is consistent with a more recent study reporting focal expression of PAX8 by the OSE (Ozcan et al, 2011). I propose

that the role of PAX8 in developmental differentiation of coelomic epithelia to endosalpingeal epithelia may by accompanied with a role in EOC development either by indicating the ectopic transport of secretory fallopian tube cells to the ovary or by being switched on in the OSE and facilitating their differentiation to Mullerian morphology as they form the inclusion cysts. The fact that some of the ovaries studied did not express PAX8 in the OSE but in the inclusion cysts may suggest that the change may be happening at the inclusion after the OSE is differentiated. The other structures that were found to express PAX8 do not originate from Mullerian or coelomic differentiation but are mesonephric type epithelia like observed in samples from the male testis and might be indicating another role for PAX8 at the developmental process where the Wolfian duct is deactivated for Mullerian development to take place. Collectively these results support that both NOSE and FTE cells should be used in models for studying EOC development.

To establish that a NOSE & FTE versus EOC mRNA differential expression assay was a good model for studying the potential significance of candidate genes in EOC development, I evaluated the expression of a handful of genes with a well-established role in EOC development. I have used the established NOSE and FTE cell lines in order to compare their expression of well-established oncogenes in EOC development with EOC cell lines. I found that the oncogenes *KRAS*, *BRAF* and *MYC* were all significantly over-expressed in EOC cell lines compared to both NOSE and FTE cell lines. A slight trend revealed that *BRAF* and *KRAS* are slightly more expressed in NOSE than FTE cell lines. This is in line with reports that the main precursor cell of low/borderline grade serous and mucinous EOC, where *KRAS* and *BRAF* mutations are mainly found, is the OSE (Cheng et al, 2005, reviewed in Naora, 2007)

An interesting observation was made regarding the differential expression of well-established tumour suppressor genes between NOSE & FTE and EOC cell lines. *BRCA1*, *BRCA2* and *TP53* although known to be implicated in EOC development as tumour suppressor genes, were found based on the NOSE, FTE and EOC expression model to have a gain of function role. This observation was supported by a similar trend of *BRCA1* and *BRCA2* being over-

expressed in SOC compared to normal FT after analysis of TCGA available mRNA data. Emerging evidence may suggest that tumour suppressor genes may function as being “oncogene-like” if exploited and activated by the cancer cell in order to surpass the stress caused by cytotoxic drugs for instance (Zheng et al, 2012). Additionally previous studies have shown that cisplatin resistant tumours have found to harbour secondary activating mutations in *BRCA1* (Swisher et al, 2008) and *BRCA1* over-expression promoted cisplatin resistance in ovarian cancer cells (Chock et al, 2010). Thus, it is conceivable since BRCA1, BRCA2 and p53 are all involved in DNA repair and survival pathways they could be rewired and used by the cancer cell to repair DNA breaks caused by cytotoxic drugs such as platinum drugs or DNA damage from endogenous replication. Another recent study has proposed a novel function of BRCA1 positively regulating the secretion levels of follistatin which is a protein found elevated in ovarian carcinoma. Knocking down *BRCA1* in an immortalised NOSE (IOSE) cell line and SKOV3, a cancer cell line that possesses wild type *BRCA1*, caused a significant decrease in the proliferation and migrating ability of the cell lines (Karve et al, 2012).

Based on the latter indication of tumour suppressor genes acting in a tumour supporting role, it may be possible that there is a limitation of using a model studying differential expression of candidate genes between NOSE & FTE and EOC lines, or even SOC and normal FT. The limitation applies to not being able to study early events in oncogenesis which would involve the loss of function for tumour suppressor genes that later are re-activated in the cancer cells to overcome stress. However, that may be not true for certain tumour suppressor genes that maintain their function in early and late stages. Specific *in vitro* transformation models where defined genetic modification will be imposed to NOSE and FTE cell lines, preferably in 3D, could better elucidate the events of early oncogenesis in EOC. Such studies have been previously performed with oncogenes such as overexpressing *KRAS* and *BRAF* in 3D IOSE cell lines (Lawrenson et al, 2011) and *HRAS* in 2D FTE cells (Shan et al, 2012, *in print*).

The difference in the expression of the investigated genes observed between the EOC and NOSE & FTE compared to SOC versus normal FT could

be explained by the differences both in the cancer and normal samples of the two models. This also reflects the well-established fact that EOC cannot be treated as a single disease since all the histological subtypes are distinct in the risk factors and molecular pathways implicated in their development. The EOC cell lines used were cell lines representing several types of EOC and the SOC in TCGA were high grade serous carcinomas. NOSE and FTE cells are both proposed to be the cell of origin for different proportions of serous and other subtypes of EOCs but each may contribute to the development of certain subtypes to a different extend. Additionally, the normal FT tissue used to generate mRNA expression data by TCGA mainly consists of stromal cells. Such high proportion of stromal cells in these samples means high proportion of genetic material of many fibroblastic and connective tissue cells which is not ideal and may bias the result of expression analyses of genes that are naturally differentially expressed in epithelial and stromal cells. For those reasons, differences in the expression patterns of candidate genes between the different models are not surprising if not indeed expected.

The EOC cell lines are originating from different subtypes of EOC and thus in following chapters I will combine the FTE and NOSE cells which to compare their expression of candidate genes with the diverse set of EOC cell lines. Another limitation of this model is that ideally I would like to investigate the expression of candidate genes between certain subtypes of EOC and specific cells of origin for each. The number of samples available and the uncertainty of histological characterisation of the EOC cell lines does not make this possible. Additionally, further work elucidating the exact cell of origin for particular types of EOC is needed so that the functional role of candidate genes can be assayed in the most appropriate model.

3.7.4 Conclusion

To conclude, this is the first study that reports establishing a large scale bank of primary NOSE cell lines and sub-culturing of primary FTE cells, rich in secretory cells to further create FTE 3D models which better reflect the *in vivo* FTE activity and architecture. I propose that studying the differential expression of genes between large numbers of NOSE & FTE and EOC cell lines is an

appropriate model for identifying candidate oncogenes involved in EOC development. The potential identification of a gain of function gene using this model however should be taken with the knowledge that the molecular events represented may not be reflecting early events of oncogenesis initiation and could very well be a tumour suppressor gene being reactivated in late stages of the disease thus found over-expressed in EOC cell lines. In following chapters I will present on how this model is used as a first line evaluation of the functional role in of candidate genes that are arising from pathway approaches and Genome Wide Association Studies (GWAS) in EOC.

Additionally, I propose that the 3D *in vitro* models of FTE cells established more closely represent the normal architecture and molecular features of the FT tissue *in vivo*. In the future, 3D models of FTE cells can be employed in the study for the early events of tumour initiation which is of critical importance given the proposal that a proportion of several subtypes of EOC arise from the epithelial mucosa of the FT. It is of critical importance to understand the molecular events that take place in the fallopian tube mucosa to make it more prone to neoplastic transformation. One hypothesis could be that the phenotype of FTE cells is influenced by the mitogenic environment of the ovarian stroma. Alternatively, the region on transition between FTE cells and ovarian mesothelial-type epithelial cells is inherently more prone to neoplastic transformation. Both FTE and NOSE 3D models could be used to impose defined genetic modifications that are characteristic for different subtypes of EOC in order to elucidate the molecular events that take place for EOC initiation and development to occur. Finally, the 3D *in vitro* models of the FTE cells can be of great value for studying developmental molecular processes during embryogenesis and also benign fallopian tube diseases such as salpingitis and pelvic inflammatory disease. Ultimately, it is hoped that such the FTE 3D *in vitro* models can give much needed insight into the biology of fallopian secretory epithelial cells and the pathogenesis of diseases associated with them. This knowledge will be invaluable in increasing our ability to diagnose and treat benign and malignant disease arising in the fallopian tubes.

♣ Publications containing work from this chapter

Lawrenson K, Sproul D, Grun B, **Notaridou** M, Benjamin E, Jacobs IJ, Dafou D, Sims AH, Gayther SA." Modelling genetic and clinical heterogeneity in epithelial ovarian cancers", Carcinogenesis. 2011 Oct;32(10):1540-9.

Lawrenson K, **Notaridou** M, Lee N, Benjamin E, Jabocs I., Jones C, Gayther SA *"In vitro* Three-dimensional Modeling of Fallopian Tube Secretory Epithelial Cells", Plos One (under review)

4 Investigating the role of common alleles in candidate susceptibility genes associated with risk and development of epithelial ovarian cancer

4.1 Introduction

Familial EOC comprises only 10% of all EOC cases and mutations in the high risk susceptibility genes *BRCA1*, *BRCA2* and members of the MMR pathway account only for 37%, 9% and 5% respectively of all hereditary ovarian cancer cases (Aarnio et al, 1999, Ramus et al, 2007, Bonadona et al, 2011). The remaining unattributed risk for EOC familial cases could be accounted for by other genes of low/moderate penetrance in the population which could also contribute to the risks for sporadic ovarian cancer in the population. The approach used for identification of low penetrance genes for EOC is genetic association studies.

Genetic association studies are aiming in identifying genetic variants (SNPs) that are associated with EOC by comparing the frequencies of the genotypes of selected SNPs in unrelated subjects with EOC compared to healthy individuals. The most common approaches used are either a candidate gene or a pathway approach by genotyping SNPs within genes with a predicted function in EOC development even though now they have been superseded by the GWAS approach where SNPs are genotyped throughout the human genome. Initial candidate gene approaches have not been very successful in identifying strong associations with risk in EOC because the candidate genes were selected based on their predicted and not known role in ovarian carcinogenesis, or because the genotyped SNPs were initially only non-synonymous SNPs in the hope of identifying the causal genetic variation within amino-acid changing polymorphisms.

Additionally, the power of individual studies has been limited due to small number of samples, a problem tackled by the collaboration of several groups

into consortium approaches. The putative causal SNP associated with risk in the development of EOC may possibly be a non-coding SNP. Association studies can identify the putative causal SNP associated with the risk, and that SNP might be a non-coding SNP (regulatory SNP). In association studies the polymorphisms typed serve as surrogates for the real causal SNP around the candidate locus and this is achieved by using tagging SNPs (tSNPs) which provide genotype information about all the SNPs that are found in LD with it.

Chromosomes that are implicated in EOC development can be identified by array CGH (aCGH) and LOH studies, and tumour suppressor genes within some of the identified loci have been identified (reviewed by Despierre et al, 2010). *BRCA1* and *BRCA2* high risk ovarian cancer genes were recorded with LOH in ovarian cancer of ~60% and 40% respectively (Sato et al, 1991, Futreal et al, 1994) with a larger more recent study reporting higher *BRCA1* and *BRCA2* LOH frequencies of 82% and 69% respectively both in hereditary and sporadic ovarian tumours (Brozek et al, 2009). The exact genetic events leading to the somatic alterations identified by LOH or CGH studies are complicated to translate without further research. Traditional mechanisms leading to LOH (or allelic imbalances) are loss of whole chromosomes or chromosomal regions. Many are random passenger changes such as translocations and then loss of a rearranged chromosome. Selection of clones with these changes could be due to mutations or epigenetic changes being the second hit and knocking out an important gene. Allelic imbalances can occur where the LOH is linked to preferential loss of a specific allele, suggesting that this SNP is functionally important in pathogenesis in the sense that is implicated in a mechanism that is responsible for loss of the encoded protein in tumours. As the copy number of genes of interest is unknown, what is recorded as LOH could be any form of allelic imbalance. Previous studies have reported LOH in genes where the second hit was not attributed to mutations, and allele specific LOH has been previously proposed as an alternative mechanism functionally contributing to disease development (Kittinyom et al, 2004). Furthermore, identifying significant LOH frequency in a chromosomal locus specifically involves particular genes in this locus. Many methods have been used to identify candidate genes such as mutational analysis and positional cloning approaches.

A popular approach used to identify the functional importance of LOH and chromosomal deletions is microcell-mediated chromosome transfer (MMCT) and is a reliable strategy to identify candidate genes with functional relevance to EOC. Briefly, ovarian cancer cells with the aberrated chromosome are fused with microcells generated from a donor normal cell line that contains single copies of the human normal chromosomes tagged with a selectable marker. MMCT allows for genetic mapping through a process of functional complementation of the tumourigenic phenotype where the resulting hybrid cell lines containing the wild-type chromosome can be tested for phenotypic changes with assays that can suggest a reversal of transformed phenotype such as anchorage independent growth assays, growth rates, tumourigenicity in nude mice, and expression of specific genes (Schulz et al, 1987, Kruzelock et al, 2000, Doherty and Fisher, 2003, Dafou et al, 2009). The genes identified by MMCT can be used in pathway approach association studies where SNPs within the candidate genes can be investigated.

Previous CGH analysis within our group has identified several chromosomes that were completely or partially deleted in chromosomes 4-6, 13-15 and 18 (Ramus et al, 2003). These chromosomes were transferred using MMCT in two EOC cell lines TOV112D and TOV21G and the donor-recipient hybrids were tested for their neoplastic phenotype *in vitro*. TOV112D^{+18} and TOV21G^{+18} hybrids exhibited neoplastic suppression in anchorage independent growth assays and invasion assays as well as reduced tumourigenicity in immunosuppressed mice compared to parental TOV112D and TOV21G, suggesting that chromosome 18 may harbour important tumour suppressor genes with functional importance in EOC development (Dafou et al, 2009).

Microarray gene expression analysis was used to identify differentially expressed genes between the hybrid and the parental cell lines for chromosome 18. More than 500 genes were identified to be up or down-regulated in the hybrid cell lines and the genes were ranked according to the P-value (cut-off: 0.05) and the fold-change difference in expression (>1.5 fold). The top 190 genes were selected and evaluated for their known predicted function, the number of common SNPs and how efficiently they could be tagged according to their sizes. Nine genes whose differential expression between

hybrid and parental cell lines was consistent between the different hybrid clones were selected based on their possible role in the development of EOC and how efficiently they could be tagged.

We hypothesized that SNPs within those genes may have an association with ovarian cancer and this was tested within our group. The candidate gene selection and SNP tagging was performed in the context of studying genetic susceptibility and survival of candidate genes in chromosome 18 and 16 (Quaye et al, 2009, Notaridou et al, 2010). Nine genes were selected from the MMCT-18 study: *AIFM2, AKTIP, AXIN2, CASP5, FILIP1L, RBBP8, RGC32, RUVBL1* and *STAG3*. These genes met the selection criteria of having a known function, exhibiting consistent fold change data between the different hybrid and parental cell lines, having between 3 and 20 tSNPs and at least one common variant (MAF>0.05) per 2kb of gene. The function of the genes was obtained from Genecards (http://genecards.org/) and their predicted role in ovarian cancer and other cancers was researched in Oncomine (http://www.oncomine.org/). Haploview software was used to sufficiently identify the tSNPs within the selected genes with SNP selection criteria MAF≥0.05 and HWE ≥0.01 using HapMap Data Release 22. The tSNPs of the selected genes were separated in multiplex panels using iPLEX Gold Assay Design software after testing several different gene combinations to create the most efficient multiplexes.

I further chose to test the functional importance of the selected SNPs by performing allele specific LOH analysis in ovarian tumour samples. My hypothesis was that associated SNPs identified may be implicated in EOC development by having a functional role in somatic alterations.

Aims of this chapter:

1. Evaluate the potential of allele specific LOH (AS LOH) for SNPs in candidate genes emerging from the MMCT-18 functional complementation study in primary ovarian tumours.
2. Evaluate overall LOH in candidate genes of the MMCT-18 study and determine the effect of LOH in patients' overall survival.
3. Evaluate differential expression of some candidate genes between NOSE and EOC cell lines

4.2 Allele specific loss of heterozygosity analysis of MMCT-18 candidate ovarian cancer moderate susceptibility alleles

The function of the selected candidate genes from the MMCT project and the number of tSNPs that were genotyped and used for AS LOH analysis is shown at Table 4.1. A total of 58 tSNPs of the 9 MMCT-18 candidate genes was selected to be genotyped in 2 iPLEX multiplex assays, a 33plex and a 27plex assay. Additionally to the tSNPs of the candidate genes, one SNP from *BRCA1* and one from *BRCA2* were included in the assays, in order to serve as controls for assaying LOH.

DNA extraction with needle microdissection of 301 ovarian tumour samples from formalin fixed paraffin embedded samples of the Malova collection of samples was performed as described in the methods section. The efficiency of the DNA extraction was checked by PCR amplification of a control fragment (*BRCA1* exon 17). Germline matching DNA was also available, extracted from the blood of the selected 301 patients.

Analysis for LOH was performed in 301 and 239 tumour DNA samples and their matching genomic DNA in a 33plex and a 27plex SNP reaction respectively. In an initial run 181 tumour DNA and their matching genomic DNA samples were genotyped for the total of 60 tSNPs. In a subsequent run 120 extra samples were genotyped for the 33plex and only 58 additional samples were genotyped for the 27plex. The reason for the difference in the amount of samples run for each set of SNP plexes was merely because of the cost and the limited availability of reagents at the time.

The ratio of the allele peak heights between the tumour and the germline DNA for heterozygous (informative) individuals was used to determine LOH as previously described in the methods section. AS LOH was assayed based on the deviation from the expected 1:1 random ratio of loss of one allele compared to the other allele. Evidence of AS LOH would be shown by preferential loss of one allele compared to the other. The P value was calculated using a two-sided Fisher's test.

Quality control criteria were applied to the data. Deviation from Hardy Weinberg Equilibrium (HWE) was assessed using the χ^2 test which measures the deviation from the expected frequencies. SNPs deviating from HWE may confound trait-allele association as they are thought to reflect genotyping error. Studies that deviated for HWE at $P<10^{-4}$ were excluded from the analysis. SNPs where the call rate of the genomic DNAs was <90% were excluded. Moreover, based on 13 duplicates of genomic DNA, SNPs with less than 98% genotype concordance between the duplicate samples were excluded from the analysis. Additional secondary QC for concordance of recording LOH was performed. Genotyping of 95 paraffin samples was repeated for the 33Plex in order to confirm that the peak heights ratios were reproducible, thus in order to validate the LOH data. Due to limited reagent availability only 13 duplicate samples were repeated for the 27plex. If SNPs were recorded as having LOH for a specific allele in the initial run and the peak height of the same allele was found to be reduced in the validation run with peak height ratios similar to the LOH values (L<0.8 or >1.4), they were classed as concordant. If the concordance of LOH between the duplicate samples was <85% the assay was excluded from the analysis. For duplicate samples that were concordant the log average of the L value was used in the analysis. A summary of the LOH concordance analysis for the selected SNPs is shown in Appendix 2, Table1.

Following all the steps of QC, out of the 60 SNPs genotyped 1 was excluded from the analysis because it was not polymorphic, 1 assay failed (failed PCR), 7 were excluded because they failed on genomic call rate. 5 were excluded because of discordances in the genotypes of genomic duplicates and 9 were excluded because of LOH discordances. A summary of the quality control analysis for the selected SNPs is shown in Appendix 1, Table 1.

Quality control analysis was also performed for the samples. Genomic and paraffin samples that failed in >20% of the assays were excluded. 9 and 10 genomic samples were removed from the 33plex and 27plex respectively at this stage. 6 and 1 paraffin samples were also removed from the 33plex and 27plex respectively. The remaining number of samples after these first steps of sample QC was 286 and 228 tumour DNA samples and their matching genomic DNA for the 33plex and a 27plex respectively. Additionally, genomic DNA plates that

demonstrated <90% call rate were excluded from the analysis. A different number of samples was removed from each SNP at this stage (Table 4.2, column “Number of samples after QC”). The final amount of samples analysed per SNP reflected the number of samples that remained after all the QC analysis as well as samples that failed for individual assays even though they had passed the overall pass rate. Additionally, samples that were discordant for LOH were removed from the analysis (Table 4.2, column (“Final number of samples analysed”).

One SNP for each of the *BRCA1* and *BRCA2* genes were included as controls for detecting LOH. These SNPs showed that LOH can be detected using the iPLEX genotyping assay, and the L value was a non-subjective measure of the difference in the allele peak height ratios. LOH for rs799917 in *BRCA1* was 56% and 28% for rs144848 in *BRCA2*. This was in consistent with the high rates of LOH already published for these genes. However *BRCA1* rs799917 and *BRCA2* rs144848 SNPs had low LOH concordance ratios, 75% and 67% respectively. One needs to bear in mind that only 4 and 3 duplicates that were informative were included for those assays respectively (part of the 27plex). One discordant result in each assay caused the concordance percentage to drop to 75% and 67% for *BRCA1* and *BRCA2* respectively (Appendix 2, Table 1).

The remaining 37 tSNPs after QC analysis were analysed for allele specific LOH. The results of analysis for allele specific LOH for all genotyped SNPs are summarised in Table 4.2. I found *STAG3* rs1637001 [A/G] SNP demonstrating allele specific LOH. Out of 190 tumours that were informative (heterozygous) for this SNP, 16 showed evidence of LOH and there was preferential loss of allele A in 94% of the cases (P=0.015, Figure 4.1). Analysis of 25 informative duplicate samples for this SNP revealed a 99% concordance for recording LOH and for all 3 duplicate samples that exhibited LOH 100% preferential loss of allele A was confirmed.

Array comparative genomic hybridisation (aCGH) analysis of ovarian tumours had been performed for the Malova set in the context of another project within our group (Dr Christopher Jones). Therefore, I investigated whether the preferential loss of allele A for rs1637001 in *STAG3* was the result of deletion of the common A allele or amplification of the rare G allele using these data. aCGH was performed for 12 of the ovarian tumours that showed allelic imbalance for rs1637001. Eight of these tumours, corresponding to cases 278, 1628, 1455, 1078, 262, 439, 3778, 1149 showed copy number gain for *STAG3* and four tumours (cases 629, 1618, 977, 1661) were copy number neutral (Figure 4.2); none of the tumours showed deletion of *STAG3*. These data suggested that, for the majority of tumours, there was amplification of the rare G allele rather than deletion of the common A allele. For tumours that were copy number neutral for *STAG3*, it is possible that the somatic alterations that were revealed by LOH analysis were not detected by aCGH due to subsequent amplification of the remaining allele.

Gene	Expression: parental vs hybrid cell lines	Cytoband	Function	Gene size (bp)	Number of tSNPs	Number of tSNPs after QC
AIFM2	3 fold up in TOV112D & TOV21G hybrids	10q22.1	TP53-induced apoptosis; overexpression induces apoptosis	34,711	13	7
AKTIP	4 fold up in TOV112D & TOV21G hybrids	16q12.2	Apoptosis; interacts with PKB/Akt;	11,978	4	4
AXIN2	5 fold up in TOV112D hybrids	17q23-q24	Inhibitor of β-catenin in Wnt signalling pathway; LOH in breast & other cancers.	33,084	8	5
CASP5	7 fold up in TOV21G hybrids	11q22.2-q22.3	Regulation of apoptosis.	14,729	8	5
FILIP1	5 fold up in TOV112D hybrids	3q12.1	Down regulated in ovarian cancer.	281,369	7	5
RBBP8	7 fold up in TOV112D hybrids	18q11.2	RB1 binding protein; transcriptional regulation of *BRCA1*; DNA repair; TSG	93,155	3	3
RGC32	3 fold down in TOV112D hybrids	13q14.11	Cell cycle progression regulation; induced by p53 in response to DNA damage.	13,323	5	3
RUVBL1	25 fold down in TOV112D & TOV21G hybrids	3q21	Interacts with *MYC*; involved in cell growth	42,857	7	3
STAG3	9 fold up in TOV21G hybrids	7q22.1	Component of cohesin complex; chromosome segregation	43,764	3	2

Table 4.1**:** Candidate genes selected from the MMCT-18 project. The predicted functions for these genes and tSNPs within them genotyped by iPLEX are presented. A number of SNPs was excluded from the analysis after quality control analysis.

Gene	SNP	Number of samples after QC	Final number of samples analysed	NI	Hets	Samples with LOH	LOH %	LOH allele 1	LOH allele 2	Ratio of allele loss	P value
AIFM2	**rs1053495**	228	228	199	29	12	41	7	5	1.4	1.000
AIFM2	**rs2394656**	228	227	168	59	14	24	7	7	1.00	1.000
AIFM2	**rs6480440**	286	270	187	83	21	25	10	11	1.1	1.000
AIFM2	**rs2280201**	228	228	185	43	12	28	6	6	1.00	1.000
AIFM2	**rs10999147**	286	286	242	44	9	20	3	6	2.0	0.637
AIFM2	**rs3750772**	228	228	194	34	5	15	3	2	1.50	1.000
AIFM2	**rs10999152**	179	172	118	54	23	43	9	14	1.56	0.554
AKTIP	**rs9931702**	228	224	125	99	48	48	27	21	1.29	0.683
AKTIP	**rs17801966**	179	176	139	37	15	41	10	5	2.00	0.462
AKTIP	**rs7189819**	227	224	139	85	23	27	10	13	1.30	0.768
AKTIP	**rs3743772**	286	259	238	21	10	48	7	3	2.3	0.650
AXIN2	**rs11868547**	228	225	115	110	58	53	27	31	1.15	0.853
AXIN2	**rs7591**	286	284	155	129	74	57	30	44	1.5	0.322
AXIN2	**rs7210356**	286	284	224	60	36	60	20	16	1.3	0.814
AXIN2	**rs4791171**	221	215	133	82	42	51	22	20	1.1	1.000
AXIN2	**rs3923086**	228	228	123	105	52	50	26	26	1.00	1.000
CASP5	**rs518604**	286	281	141	140	29	21	14	15	1.1	1.000
CASP5	**rs3181328**	228	228	188	40	10	25	3	7	2.33	0.650
CASP5	**rs9651713**	228	222	185	37	12	32	3	9	3.00	0.400
CASP5	**rs3181175**	173	172	119	53	13	25	8	5	1.6	0.695
CASP5	**rs2282657**	228	147	70	77	24	31	13	11	1.2	1.000
FILIP1L	**rs796977**	286	280	164	116	18	16	8	10	1.3	1.000
FILIP1L	**rs793477**	179	177	146	31	7	23	5	2	2.50	0.576
FILIP1L	**rs9864437**	286	285	193	92	24	26	9	15	1.7	0.561
FILIP1L	**rs6788750**	228	223	118	105	23	22	7	16	2.29	0.231
FILIP1L	**rs12494994**	286	266	193	73	24	33	14	10	1.4	0.772
RBBP8	**rs7239066**	228	228	182	46	13	28	7	6	1.17	1.000
RBBP8	**rs11082221**	228	228	212	16	6	38	2	4	2.00	1.000
RBBP8	**rs9304261**	138	137	88	49	20	41	11	9	1.2	1.000
RGC32	**rs11618371**	286	286	228	58	25	43	17	8	2.1	0.252
RGC32	**rs995845**	286	273	167	106	36	34	13	23	1.8	0.341
RGC32	**rs975590**	228	226	145	81	23	28	12	11	1.09	1.000
RUVBL1	**rs3732402**	286	283	143	140	33	24	14	19	1.4	0.554
RUVBL1	**rs4857836**	286	286	173	113	21	19	17	4	4.3	0.052
RUVBL1	**rs9821568**	228	223	166	57	10	18	7	3	2.33	0.650
STAG3	**rs11762932**	286	283	195	88	13	15	6	7	0.9	0.523
STAG3	**rs1637001**	286	282	190	92	16	17	15	1	15.0	**0.015**

Table 4.2: LOH and allele specific LOH in tumours for 37 tSNPs across the MMCT-18 candidate genes. NI: Non informative (homozygous), Hets: heterozygotes.

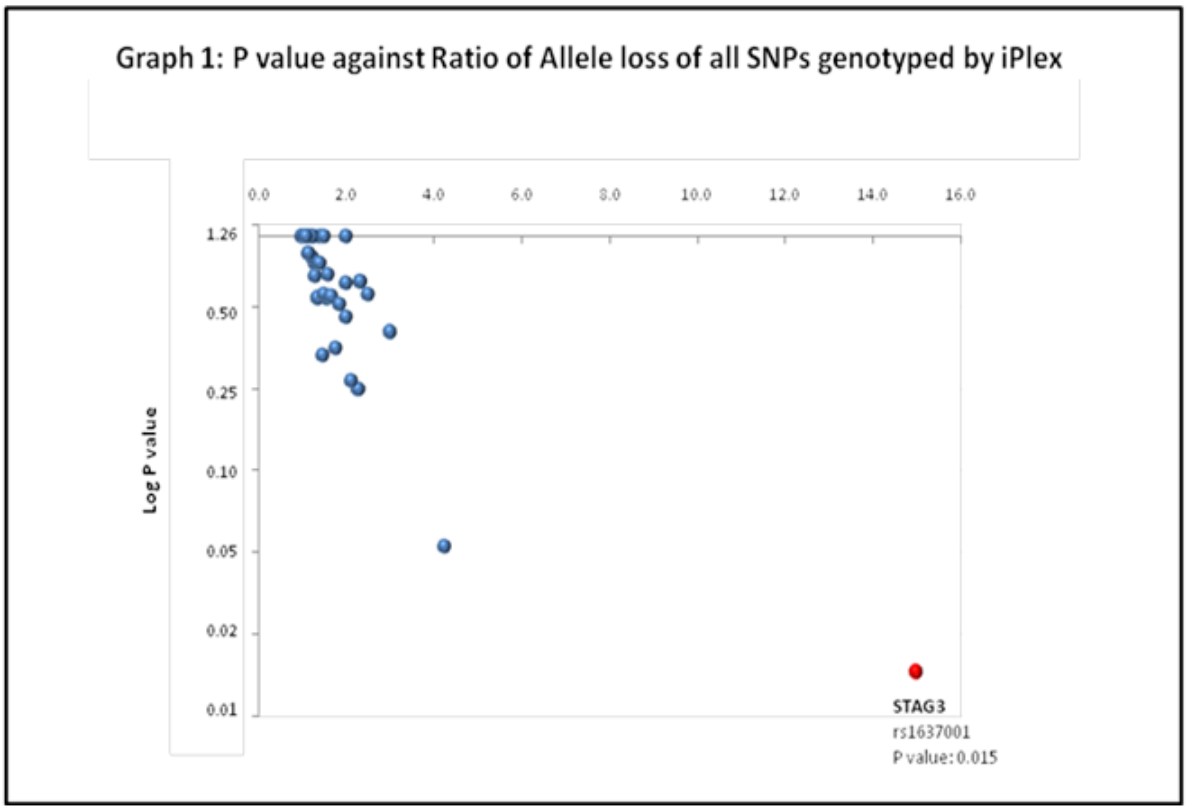


Figure 4.1**:** P values against ratio of allele loss of MMCT-18 tSNPs genotyped by iPLEX. SNP rs1627001 [A/G] of *STAG3* demonstrated 94% preferential loss of the common allele A and is highlighted on the graph in red.

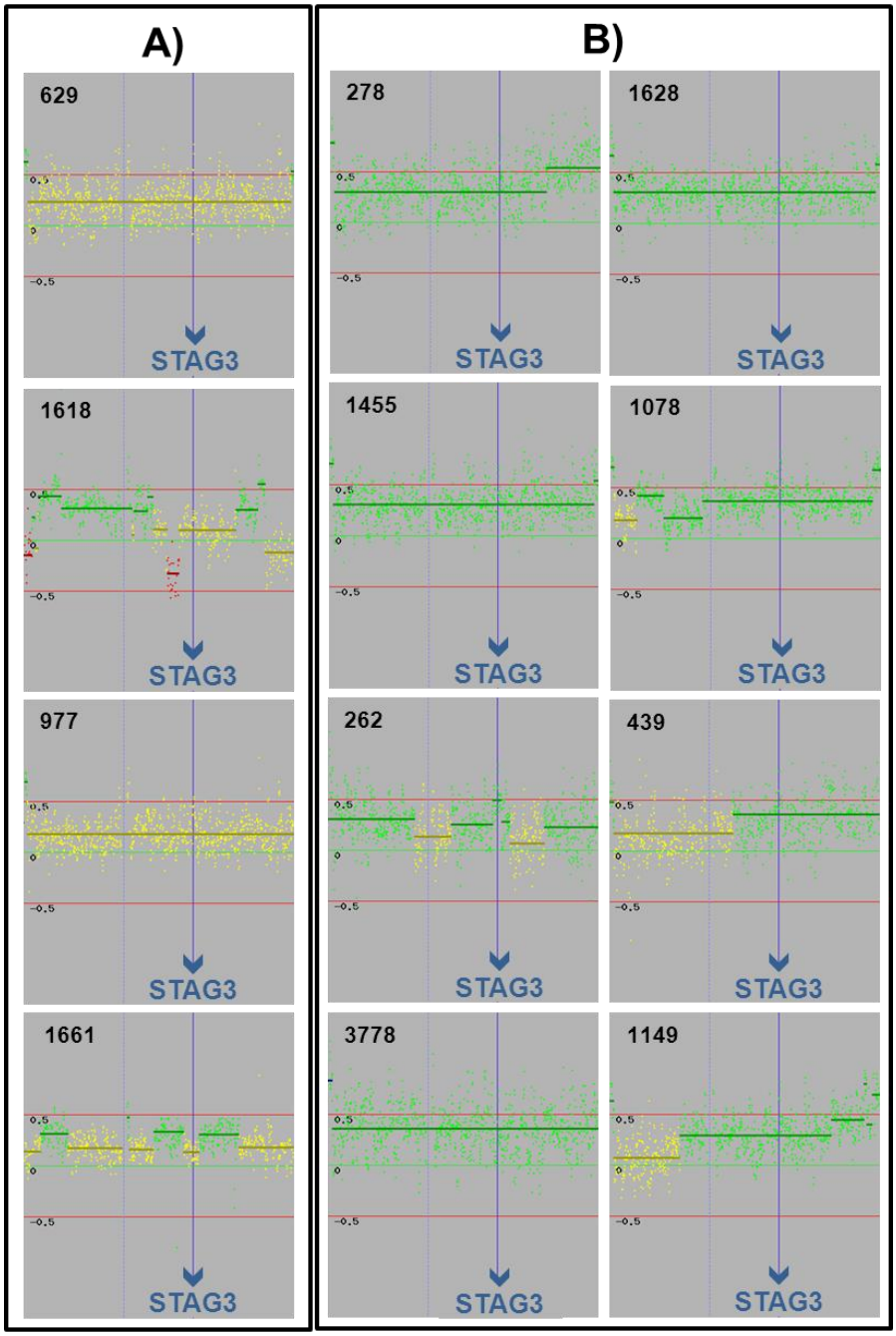


Figure 4.2**:** aCGH profiles of 12 of the 16 tumours that exhibited LOH for rs1637001 of *STAG3*. Regions around the zero baseline are highlighted in yellow and indicate no change in chromosomal copy number, regions significantly below the baseline are highlighted in red and indicate loss. Gain or amplification is indicated by green in regions that are found significantly above the baseline. The probability of having gain or loss was calculated by a log2 ratio number and the cut-off indicating loss or gain was distinct for each aCGH experiment. The figures presented are a magnification of the aCGH profiles specifically for chromosome 7 and *STAG3* is mapped on each figure. A) Tumours that showed neutral copy number, B) tumours that showed copy number gain.

4.3 Evaluating LOH frequency across the MMCT-18 candidate genes

LOH of all tSNPs within the genes is demonstrated in Figure 4.3 and individual LOH per SNP shown in Table 4.2. The role of the candidate genes in ovarian cancer was evaluated by assaying for overall LOH within the gene. Every patient sample was genotyped for tSNPs within a gene, and if a patient demonstrated LOH in at least one of the tSNPs within a gene, it was counted as LOH observed for this sample for the candidate gene. The results of the analysis for overall LOH across the candidate genes are shown on Table 4.4. The highest LOH was observed in *AXIN2* which demonstrated 64% LOH, *AKTIP* with 46% LOH and *RGC32* with 36% LOH.

I also looked at LOH across the candidate genes stratified by subtype. The number of samples available was small so I grouped the subtypes and compared the serous with the ECs, MCs, CCCs and OthCs combined. LOH was significantly higher in serous samples compared to the rest subtypes in *RBBP8* gene with a $P<10^{-4}$. No significant differences were seen for the LOH in other genes between the different subtypes.

Histology	All samples	AIFM2 LOH	AKTIP LOH	AXIN2 LOH	CASP5 LOH	FILIP1 LOH	RBBP8 LOH	RGC32 LOH	RUVBL1 LOH	STAG3 LOH
SC	151	36	38	89	43	36	28	37	27	14
CCC, EC, MC, OthC	135	19	25	61	27	27	6	26	24	16
P value	1	0.122	0.329	0.224	0.229	0.5781	**0.0009**	0.4062	1	0.5679

Table 4.3: LOH % across MMCT-18 candidate genes in ovarian tumours stratified by histological subtype. SC: Serous carcinoma, MC: Mucinous Carcinoma, OthC: Other Carcinoma, CCC: Clear Cell Carcinoma, EC: Endometrioid Carcinoma. The P value is indicating statistical significance of difference of LOH occurrence between different histological subtypes and is calculated using Fisher's exact test.

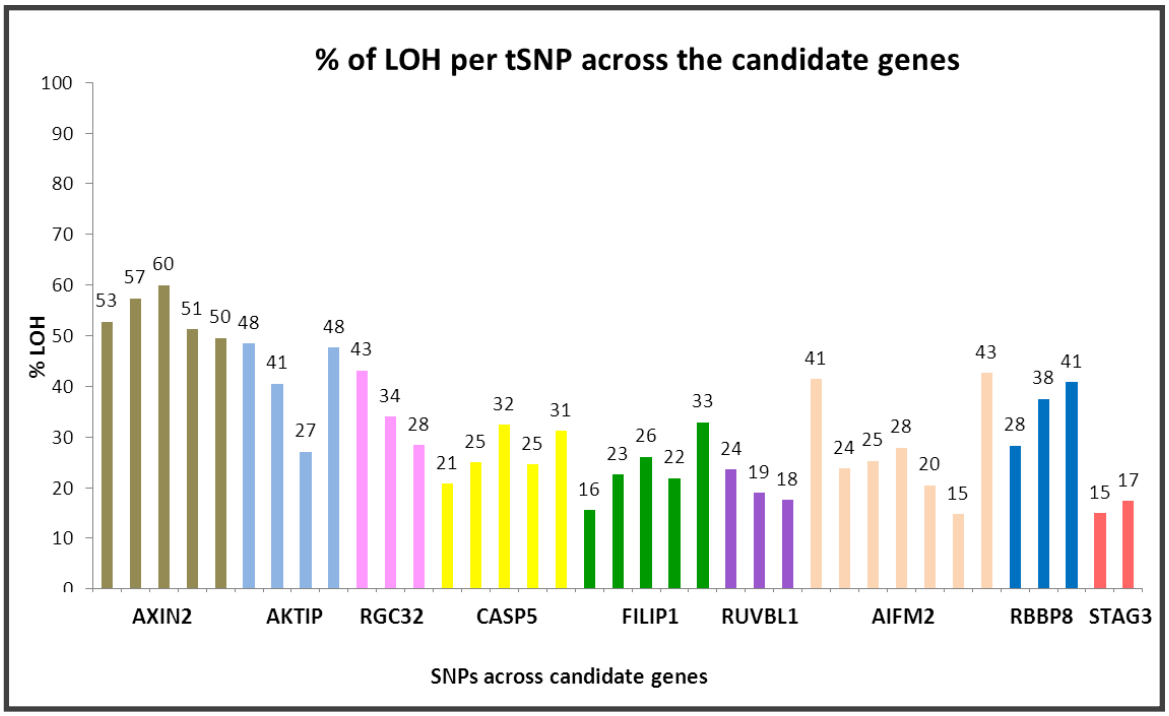


Figure 4.3: Histogram demonstrating LOH for all tSNPs within the MMCT-18 candidate genes.

Gene	Number of SNPs genotyped	Heterozygotes	Samples with LOH	% LOH
AXIN2	5	235	150	64
AKTIP	4	137	63	46
RGC32	3	185	63	34
CASP5	5	214	70	33
FILIP1L	5	196	63	32
RUVBL1	3	167	51	31
AIFM2	7	202	55	27
RBBP8	3	148	34	23
STAG3	2	162	30	19
BRCA1	**1**	97	54	56
BRCA2	**1**	89	25	28

Table 4.4**:** Summary of the LOH% identified for the MMCT-18 candidate genes. A sample was considered to have LOH within a gene if at least one of the tSNPs was exhibiting LOH.

Finally, I decided to examine whether LOH in the candidate genes correlated with loss found by aCGH analysis. Based on the genomic positions of the candidate genes indicated by Build 36, the middle of each gene was mapped on the aCGH profiles of the samples used to investigate LOH in order to find whether the corresponding region had loss, copy number neutral or gain/amplification. No chromosome segmenting was detected within the positions of the genes of interest indicating that mapping the middle of the gene was a reliable representation of the copy number status of the region. Only 95 of the 286 samples in this study had aCGH data available. Therefore only a small number of samples with LOH had available aCGH data. The aCGH result was compared with the LOH recorded for each candidate gene (Table 4.5).

I compared the percentage of samples with LOH with the percentage of samples with loss found by aCGH. I found that LOH in genes *AKTIP*, *CASP5* and *RGC32* was consistent with the percentage of tumours that were found to have loss in aCGH analysis. LOH in genes *AXIN2* and *FILIP1L* did not correlate with copy number variation found by aCGH which revealed copy number neutral in the mapped regions for the majority of the samples. This could be the result of deletion being masked by a subsequent amplification and therefore not detected as loss. Another explanation could be because LOH is caused by microdeletions that cannot be detected by aCGH which is reproducible in detecting changes in regions >0.5 million basepairs.

AIFM2, *RUVBL1, RBBP8* and *STAG3* exhibited higher percentage of samples with gain and the number of samples with loss did not correlate with the percentage of LOH found for these genes. Comparison of the number of samples that exhibited LOH per gene with the aCGH profile revealed similar conclusions, with several of the samples with LOH showing not only loss but also no changes or gain in the copy number by aCGH (Table 4.6). LOH across genes and individual samples could likely be due to amplification of the remaining allele rather than deletion of the lost one.

Gene	Position (Build 37)	% of samples with loss	% of samples with no change	% of samples with gain	% LOH across gene
AIFM2	chr10:71872030-71892690	2.9	69.1	27.9	27.2
AKTIP	chr16:53525192-53537170	29.4	63.2	7.4	46.0
AXIN2	chr17:63524681-63557740	7.4	85.3	7.4	63.8
CASP5	chr11:104864967-104893895	21.2	68.2	10.6	32.7
FILIP1L	chr3:99551988-99833349	3.0	89.4	7.6	32.1
RBBP8	chr18:20513295-20606449	6.0	76.1	17.9	23.0
RGC32	chr13:42031542-42045013	29.4	67.6	2.9	34.0
RUVBL1	chr3:127799800-127842671	3.2	65.1	31.7	30.5
STAG3	chr7:99775538-99812010	1.5	42.4	56.1	18.5

Table 4.5: aCGH analysis indicating the % of tumours found to have copy number changes for the MMCT-18 candidate genes. Copy number was evaluated after mapping in the middle of each candidate gene. Percentage of LOH across candidate genes is also shown to compare. Highlighted in green are genes that LOH correlates with loss, in blue are genes where LOH does not correlate with loss and the majority of the samples have no CN change, and in orange genes where LOH does not correlate with loss as the majority of samples are showing gain in the aCGH analysis.

Gene	Number of samples with LOH	Number of samples with loss	Number of samples with no change	Number of samples with gain
AIFM2	20	3	15	2
AKTIP	23	9	13	1
AXIN2	57	8	45	4
CASP5	28	3	22	3
FILIP1L	19	0	17	2
RBBP8	13	1	11	1
RGC32	32	5	23	4
RUVBL1	23	1	13	9
STAG3	17	0	5	12

Table 4.6: Number of tumours exhibiting copy number changes for the MMCT-18 candidate genes. Tabulated is the number of samples found to have copy number neutral, loss or gain evaluated by aCGH analysis mapped at the middle of each candidate gene.

4.4 The association between LOH status and ovarian cancer survival

I decided to investigate if the candidate genes had any involvement in patient survival from ovarian cancer and thus I evaluated survival of the Malova cases showing LOH in the primary tumours for the 9 candidate genes.

The patients from the MALOVA tumour set used for the LOH analysis have been recruited in a study over a 10 year period and follow up data were taken to assess the number of survivors and those who die over that period of time. The start of the survival period was defined at the date blood was taken from each patient and the end was either death or end of the 10 year follow up period for the study. Survival was measured by the hazard ratio (HR) which is the conditional probability of death. Similarly with survival analyses that use genetic polymorphisms or gene mutations as variables that can affect survival, in the current study LOH and no-LOH were used as the variables that could affect survival. If LOH in a particular gene was significantly associated with survival then a statistically significant difference of the HR would be observed between the patients with LOH and patients with no-LOH in that gene.

Similar to the analysis of overall LOH across genes, LOH for a case was recorded if at least one informative (heterozygous) tSNP in a gene showed LOH. The survival analysis was performed in STATA statistical software using Cox proportional hazards regression model to measure the effect of LOH of the MMCT-18 candidate genes over the time death takes to occur within the specified 10 year period (survival). This model determined the estimates of hazard ratios (HR) for the LOH versus no-LOH groups. To illustrate the survival of the LOH versus the no-LOH groups, over a period of time, Kaplan-Meir survival curves were generated. These are graphs of time versus survival probability and are a series of horizontal steps of declining magnitude which for large enough samples sizes are representative of the true survival function for that population. The Kaplan-Meir survival curves would be statistically different from one another if the variables had a significantly differential effect on the outcome of a patient. If the time of the blood draw is different than the time of diagnosis then left truncation should be taken into account. Thus, left truncation is accounted when there is a proportion of patients with delayed entry in the

study possibly biasing the outcome in the end of the 10 year follow up. All the patients in this study entered the study at the time of diagnosis thus left truncation did not apply.

4.4.1 Univariate survival analysis

A summary of the results is shown on Table 4.7. Kaplan-Meier Curves of the significant results were also constructed (Figure 4.4). Four genes showed association of LOH with patient survival. I found that patients with LOH in the genes *CASP5* and *AIFM2* had significantly worse survival than cases without LOH in all subtypes of EOC HR=1.47 (1.04-2.08), P=0.031 and HR=1.59 (1.09-2.30), P=0.015 respectively. Cox regression analysis by histological subtype, serous versus non-serous, did not reveal an association of poor survival in LOH patients linked with a distinct subtype. Significantly worse survival was also found in cases with LOH compared with the ones with no LOH in the genes *AXIN2* and *RBBP8* when all subtypes were analysed, with HR=2.18 (1.52-3.12), P<0.0001 and HR=2.48 (1.60-3.84), P<0.0001 respectively. Analysis for survival in serous and non-serous disease revealed that the association was driven by LOH association with survival from non-serous disease in *AXIN2* and serous disease for *RBBP8*. Interestingly, LOH in *BRCA1* was not associated with worse survival when all subtypes were analysed, however, after subtype specific analysis, LOH in *BRCA1* was associated with worse survival in non-SC patients [HR=3.40 (1.38-8.48), P=0.008] but had no effect in the survival of in SC patients.

Multiple testing for the univariate analysis was performed using Bonferroni's correction (n=33, P^B<0.002: significant cut-off). The association of *AXIN2* LOH with worse survival in non-SCs remained significant. The associations of *AXIN2* and *RBBP8* LOH with worse survival in all EOC subtypes also remained significant after Bonferroni correction. However, it is puzzling why they are stronger associations than the ones observed in the distinct subtypes since they are driven only by non-serous and serous for the *AXIN2* and *RBBP8* respectively.

4.4.2 Assessing the effect of clinical factors on the survival of EOC patients

Clinical factors such as age at diagnosis, histological subtype, stage and grade of the tumour are known to affect the survival for EOC patients. These factors are known as prognostic factors as they can be used to estimate the patient's outcome and chances for survival. The survival analysis initially performed is termed as univariate as it did not take into account any effects of the clinical prognostic factors on the survival of the patients studied. Therefore, any positive associations may have been falsely created and also real associations may have been hidden due to the effect of the clinical prognostic factors.

On that note, cox regression analysis was used to estimate the effects of these factors on the 286 samples used for the analysis, using the <40 age group, serous histology, grade1 and localised/early stage as calibrators/references to compare the rest, for each prognostic factor respectively. The analysis was performed in STATA. The results of the effect of the clinical prognostic factors on EOC survival analysis in the dataset of the 286 patients are summarised in Table 4.8 and Kaplan-Meier curves of the significant results are shown in Figure 4.5. This analysis revealed that the survival of patients that were over 60 years old was significantly worse compared with the patients under 40 years old (HR=3.66 [1.15 -11.7], P=0.28). Patients within the age group between 40 and 60 years old did not show any significant survival difference to the calibrator group. Additionally, tumours of advanced/late stage (Stage III&IV) had a dramatically statistically significant increased mortality when compared with the localised tumours (Stage I &II, HR= 6.76 [4.32-10.59], P<0.0001). The histological subtype and grade of the tumour was not significantly associated with survival of the 286 patients.

4.4.3 Multivariate survival analysis

The results of all genes were adjusted for the prognostic factors age at diagnosis >60 and advanced stage, which were both significantly associated with worse survival. This kind of analysis is termed as multivariate analysis. Cox

regression multivariate analysis was performed in STATA to assess the effect of LOH of each gene on survival in all subtypes, serous only and non-serous only subtypes adjusted for prognostic factors where appropriate. The results of the multivariate survival analyses for all the genes are summarized in Table 4.7 (Multivariate analysis *). Following multivariate analysis only LOH in *RBBP8* remained associated with worse survival in all EOC subtype patients after adjusting for age and stage (HR=1.73 (1.11-2.70), P=0.015). However, following multiple testing correction the result did not remain significant.

The sample set used for this study was not a very large one and the adjustments for prognostic factors were based on their association to survival from EOC in the 286 samples alone. Dr Chris Jones, in the context of drawing a survival model for another study within our group, combining only serous cases from >400 cases of the Malova and other tumour study cohorts has approached the analysis in a different way. In this analysis Cox proportional data were generated by extracting the survival probability for a median age and then adjusted each patient's survival outcome for age and stage and residual disease. This multivariate analysis adjustment based on large numbers of tumours is a more accurate adjustment (Chris Jones, personal communication). Additionally to the multivariate analysis described before, I have also used these adjusted survival probabilities as an alternative to the previously performed multivariate survival analysis of the serous cases for LOH in each of the candidate genes. I subsequently compared the HR and P values of the two multivariate analysis for the serous tumours and found LOH in none of the selected genes LOH was significantly associated with survival outcome (Table 4.7, Multivariate analysis**).

Multivariate survival analysis when adjusting with the appropriate prognostic factors removed most significant associations found between LOH and survival with the univariate analysis, especially after identifying a very significant association of EOC stage and worse survival. I decided to look at the association of LOH with the different variables: age, stage and grade in an attempt to clarify whether LOH was associated with any of them. A Pearson chi-squared nonparametric test was performed using STATA statistical analysis software. I found that LOH in *AXIN2*, *RGC32* and *BRCA1* was significantly

associated with more advanced EOC stage (P<0.0001, P=0.016 and P=0.011 respectively) and higher grade (P=0.001, P=0.001 and P=0.01 respectively) (Table 4.9). Additionally, LOH in *RBBP8* was significantly associated with higher stage (P<0.0001), LOH in *AKTIP* with higher grade (P=0.002) (Table 4.9). Therefore, LOH in both *AXIN2* and *RBBP8* were the most significantly associated with increased EOC stage. Interestingly, LOH in *AXIN2* and *RBBP8* were extremely significantly associated with survival in all EOC subtypes in the univariate analysis before adjusting for stage. The association of LOH with stage might be responsible either for driving the positive association in the univariate analysis or by masking a positive result in the multivariate analysis.

			Univariate		Multivariate *		Multivariate **	
Gene	**Histology**	**Informative Cases**	**HR (95% CI)**	**P value**	**HR (95% CI)**	**P value**	**HR (95% CI)**	**P value**
AIFM2	**All subtybes**	202	**1.59 (1.09-2.30)**	**0.015**	1.37 (0.95-1.99)	0.095	N/A	N/A
	SC	105	1.46 (0.93-2.31)	0.101	1.41 (0.90-2.23)	0.134	2.49 (0.90-6.86)	0.079
	ECC, MC, CCC, OthC	97	1.46(0.74 -2.86)	0.271	1.29 (0.66-2.53)	0.463	N/A	N/A
AKTIP	All subtybes	137	1.21 (0.81-1.8)	0.352	1.15 (0.77-1.72)	0.494	N/A	N/A
	SC	83	1.15 (0.72-1.85)	0.563	1.17 (0.73-1.87)	0.524	0.45 (0.09-2.34)	0.345
	ECC, MC, CCC, OthC	54	1.42 (0.67-3.02)	0.366	1.15 (0.54-2.45)	0.727	N/A	N/A
AXIN2	All subtybes	235	2.18 (1.52-3.12)	**<0.0001**	1.28 (0.88-1.88)	0.19	N/A	N/A
	SC	125	1.57 (0.99-2.48)	0.056	1.32 (0.83-2.10)	0.246	2.02 (0.58-7.02)	0.27
	ECC, MC, CCC, OthC	110	**2.75 (1.54-4.93)**	**0.001**	1.48 (0.80-2.72)	0.21	N/A	N/A
RGC32	All subtybes	185	1.31 (0.91-1.89)	0.149	0.91 (0.63-1.31)	0.614	N/A	N/A
	SC	95	1.14 (0.72-1.81)	0.58	1.15 (0.73-1.83)	0.551	0.97 (0.37-2.54)	0.944
	ECC, MC, CCC, OthC	90	1.42 (0.79-2.59)	0.239	0.63 (0.34-1.19)	0.157	N/A	N/A
CASP5	**All subtybes**	214	**1.47 (1.04-2.08)**	**0.031**	1.71 (0.82-1.67)	0.378	N/A	N/A
	SC	116	1.28 (0.04-1.95)	0.251	1.18 (0.78-1.80)	0.43	2.38 (0.94-6.03)	0.068
	ECC, MC, CCC, OthC	98	1.52 (0.82-2.81)	0.187	1.43 (0.77-2.65)	0.261	N/A	N/A
FILIP1L	All subtybes	196	1.27 (0.88-1.82)	0.203	1.16 (0.80-1.67)	0.436	N/A	N/A
	SC	104	1.16 (0.74-1.84)	0.522	1.01 (0.64-1.60)	0.969	6.3 (0.27-2.20)	0.77
	ECC, MC, CCC, OthC	92	1.41 (0.77-2.58)	0.26	1.84 (0.99-3.43)	0.053	N/A	N/A
RUVBL1	All subtybes	167	0.97 (0.65-1.46)	0.886	1.13 (0.76-1.72)	0.525	N/A	N/A
	SC	94	0.88 (0.52-1.46)	0.611	0.90 (0.54-1.51)	0.697	1.40 (0.41-4.80)	0.588
	ECC, MC, CCC, OthC	73	1.2 (0.62-2.40)	0.561	1.72 (0.87-3.41)	0.122	N/A	N/A
RBBP8	All subtybes	148	2.48 (1.60-3.84)	**<0.0001**	**1.73 (1.11-2.70)**	**0.015**	N/A	N/A
	SC	84	**2.09 (1.25-3.50)**	**0.005**	1.56 (0.93-2.61)	0.091	1.78 (0.60-5.29)	0.301
	ECC, MC, CCC, OthC	64	2.53 (0.97-6.60)	0.059	2.47 (0.93-6.55)	0.07	N/A	N/A
STAG3	All subtybes	162	0.76 (0.46-1.27)	0.294	0.79 (0.48-1.32)	0.386	N/A	N/A
	SC	85	0.78 (0.40-1.54)	0.477	0.80 (0.41-1.56)	0.507	1.69E-16	1
	ECC, MC, CCC, OthC	77	0.79 (0.37-1.71)	0.552	0.81 (0.37-1.76)	0.593	N/A	N/A
BRCA1	All subtybes	95	1.60 (0.98-2.60)	0.062	0.87 (0.52-1.44)	0.581	N/A	N/A
	SC	58	0.65 (0.36-1.19)	0.16	0.62 (0.34-1.15)	0.132	0.94 (0.24-3.77)	0.935
	ECC, MC, CCC, OthC	37	**3.40 (1.38-8.48)**	**0.008**	0.85 (0.33-2.18)	0.731		
BRCA2	All subtybes	87	1.58 (0.77-3.22)	0.21	0.96 (0.45-2.04)	0.919		
	SC	57	1.34 (0.64-2.79)	0.441	1.32 (0.63-2.77)	0.459	0.48 (0.06-3.76)	0.482
	ECC, MC, CCC, OthC	30	No cases showed LOH		No cases showed LOH			

Table 4.7: Results of univariate and multivariate survival analysis relative to LOH of the MMCT-18 candidate genes. P values indicate the statistical significance of difference in survival of two groups, one exhibiting LOH across the candidate gene and the other not (*adjusted for age and/or stage where appropriate, ** analysis performed with adjusted according to median age probability adjusted for age, stage and residual disease). HR: Hazard rate is indicating the conditional probability of the death providing the patient has survived up to a period. SC: Serous Carcinomas, MC: Mucinous Carcinomas, EC: Endometrioid Carcinomas, CCC: Clear Cell Carcinomas, OthC: Other Carcinomas. N/A: No data available as this analysis was performed only in combined serous cases.

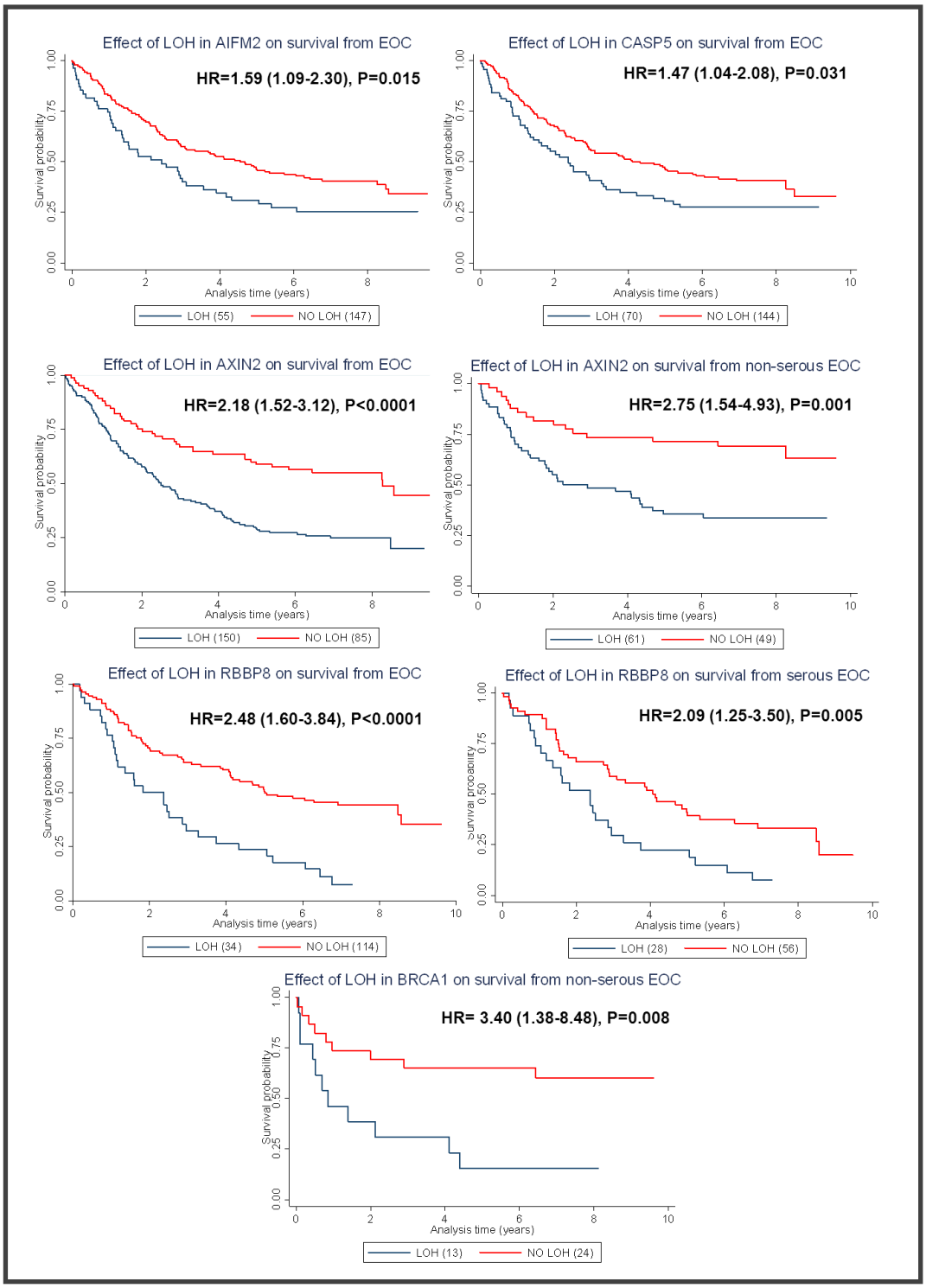


Figure 4.4: Kaplan Meier plots generated for the genes where LOH was significantly associated with survival of patients (Univariate analysis). The y axis shows the survival probability and the x axis the analysis time in years. The association of LOH in *CASP5* and *AIFM2* with worse survival was found in all subtypes. The association of LOH in *AXIN2* and *BRCA1* with worse survival was linked to non-serous subtypes. The association of LOH in *RBBP8* with worse survival was driven by serous subtypes.

	Prognostic Factor	Cases	HR (95% CI)	P value
Age at diagnosis	< 40	8	Callibrator	
	40-49	48	2.32 (0.7 -7.67)	0.169
	50-59	85	2.70 (0.84 -8.73)	0.097
	≥ 60	145	**3.66 (1.15 -11.7)**	**0.028**
Histology	Serous	151	Callibrator	
	Mucinous	37	1.15 (0.69-1.91)	0.591
	Endometriod	42	0.61 (0.37-1.02)	0.057
	Clear Cell	29	0.90 (0.47-1.72)	0.745
	Other	27	1.22 (0.75-1.98)	0.414
Tumour grade	1	64	Callibrator	
	2	104	1.07 (0.7-1.63)	0.768
	3	118	1.35 (0.9-2.03)	0.146
Tumour Stage	Localised	103	Callibrator	
	Advanced	183	**6.76 (4.32-10.59)**	**<0.0001**

Table 4.8: Summary of Cox regression survival analysis evaluating the effect of clinical prognostic factors on EOC survival. The results are based on survival data from 286 patients. Grade 1, 2 and 3 indicate well differentiated, moderately differentiated and poorly differentiated tumours respectively.

Gene		LOH- Stage	LOH- Age	LOH- Grade
AIFM2	chi2	2.309	1.049	2.701
	P value	0.129	0.789	0.259
AKTIP	chi2	1.31	4.87	12.11
	P value	0.252	0.182	0.002
AXIN2	chi2	23.198	2.716	13.984
	P value	<0.0001	0.438	0.001
RGC32	chi2	5.838	2.144	14.174
	P value	0.016	0.543	0.001
CASP5	chi2	2.1365	3.938	6.807
	P value	0.144	0.14	0.078
FILIP1L	chi2	2.956	5.519	3.01
	P value	0.086	0.138	0.222
RUVBL1	chi2	1.064	2.251	4.625
	P value	0.302	0.522	0.099
RBBP8	chi2	12.844	0.49	8.095
	P value	<0.0001	0.921	0.017
STAG3	chi2	0.374	3.227	5.169
	P value	0.541	0.358	0.075
BRCA1	chi2	6.501	1.604	9.201
	P value	0.011	0.659	0.01
BRCA2	chi2	2.132	3.8339	6.151
	P value	0.144	0.28	0.046

Table 4.9: Pearson Chi squared non parametric test revealing the association of LOH in candidate genes with stage, age and grade for the MMCT-18 candidate genes. Stage (1 degree of freedom), age (3 degrees of freedom) and grade (2 degrees of freedom).

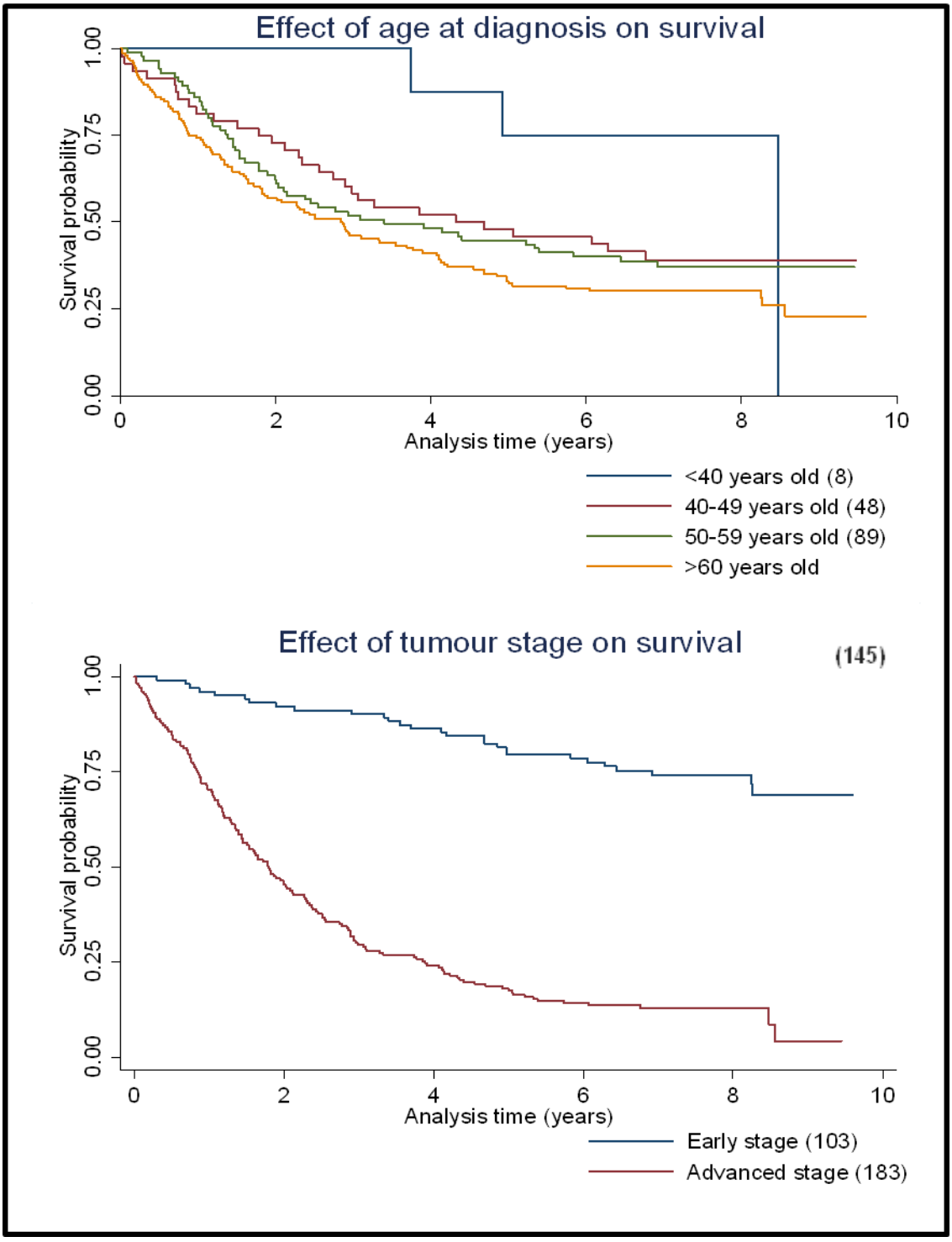


Figure 4.5: Kaplan Meier curves illustrating the effect of age and tumour stage on the survival of 286 EOC patients.

4.5 Evaluating differential expression of two MMCT-18 candidate genes between NOSE and EOC cell lines

I evaluated the differential expression of *RBBP8* and *STAG3* in a panel of 45 NOSE compared to 24 EOC cell lines normalised to two endogenous genes, GAPDH and β-actin using Taqman Real Time PCR. I chose to evaluate the expression of *STAG3* to investigate whether the allele specific LOH observed in SNP rs1637001 is functionally important in a regulatory manner affecting the expression of the gene. *RBBP8* was chosen because LOH across the gene was significantly associated with worse survival from EOC even after multivariate analysis. The reason for not investigating the differential expression for all 9 candidate genes was due to limited reagent availability and cost at the time and attention to only the most interest results was decided.

The mRNA expression data were analysed using the comparative ΔΔt method and the expression values of all cell lines were generated relative to either the lowest or highest expression of a NOSE cell line for each cell line normalized against GAPDH and β-actin. Differences in the relative expression of each candidate gene between EOC cell lines and NOSE cell lines were assessed using the nonparametric two sided Wilcoxon Rank sum test using R software and P values were generated. Statistically significant values were considered if $P<0.05$.

I have also investigated the expression of all the MMCT-19 candidate genes between high grade serous ovarian carcinomas (SOC) and normal samples (from whole FT tissue) using mRNA expression data publicly available from the TCGA. Additionally, I have investigated whether there is a correlation of expression of these genes with copy number variation found in the SOC (www.cbioportal.org.uk).

I found that expression of *RBPP8* was significantly higher in the NOSE cell lines compared to the EOC cell lines ($P=0.044$, Figure 4.7 A), which indicates a loss of function role for *RBBP8* in EOC development consistent with the increased expression of *RBBP8* in the hybrid cell lines of the MMCT-18 project. According to TCGA data there was no difference in the expression of *RBBP8* between SOC and normal samples (Figure 4.7 B) and no significant copy number change was associated with the expression of the gene (Figure

4.7 C). Comparison of *STAG3* expression between EOC and normal cell lines showed no differential expression (Figure 4.7 D) whereas TCGA data revealed that SOC had significantly reduced expression compared to normal tissue (P= 7.1×10^{-6}, Figure 4.7 E). An interesting observation is that the over-expression of the gene in SOC was characterized by copy number variation compared to normal. All the SOC samples had copy number loss, gain or amplification leading to *STAG3* over-expression indicating different mechanisms that may lead to *STAG3* under-expression in SOC compared to normal (Figure 4.7 F).

I have also looked at the differential expression of the remaining 7 candidate genes between SOC and normal FT samples and found that *AIFM2* was significantly over-expressed in SOC compared to normal (P=0.009, Figure 4.6 A). The over-expression of AIFM2 in SOC was observed in the SOC that exhibit higher copy number that the normal (Figure 4.6 B). None of the other 6 genes were differentially expressed between SOC and normal according to TCGA.

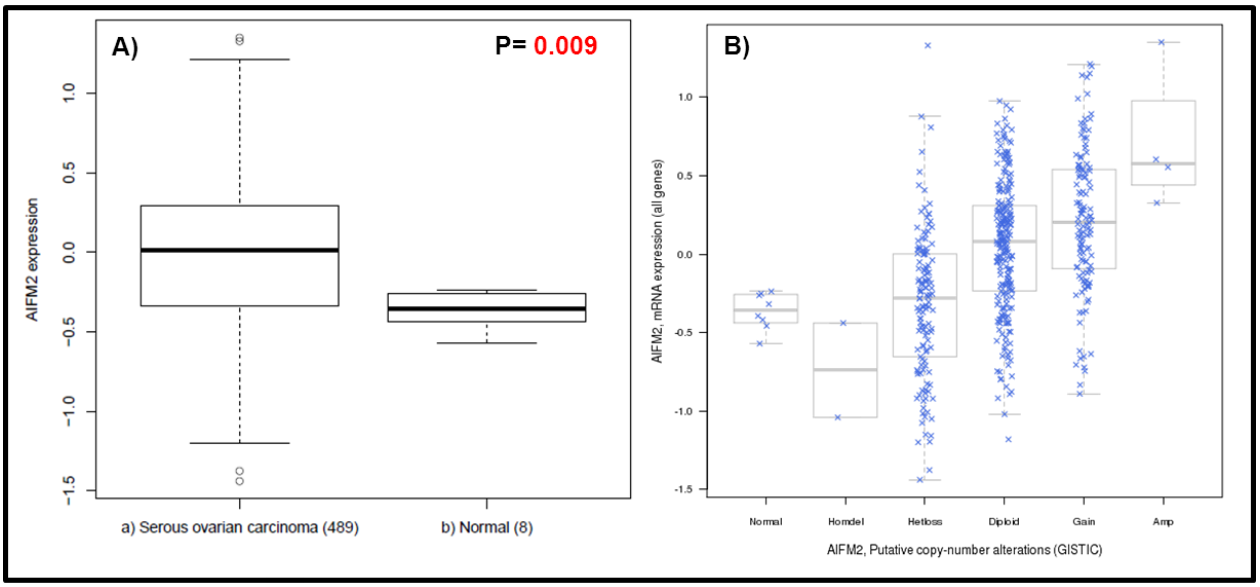


Figure 4.6: Differential expression of *AIFM2* MMCT-18 gene between SOC and normal FT tissue. A) Differential expression between SOC and normal FT for *AIFM2*. Differential expression relative to copy number variation between SOC and normal FT for *AIFM2.*

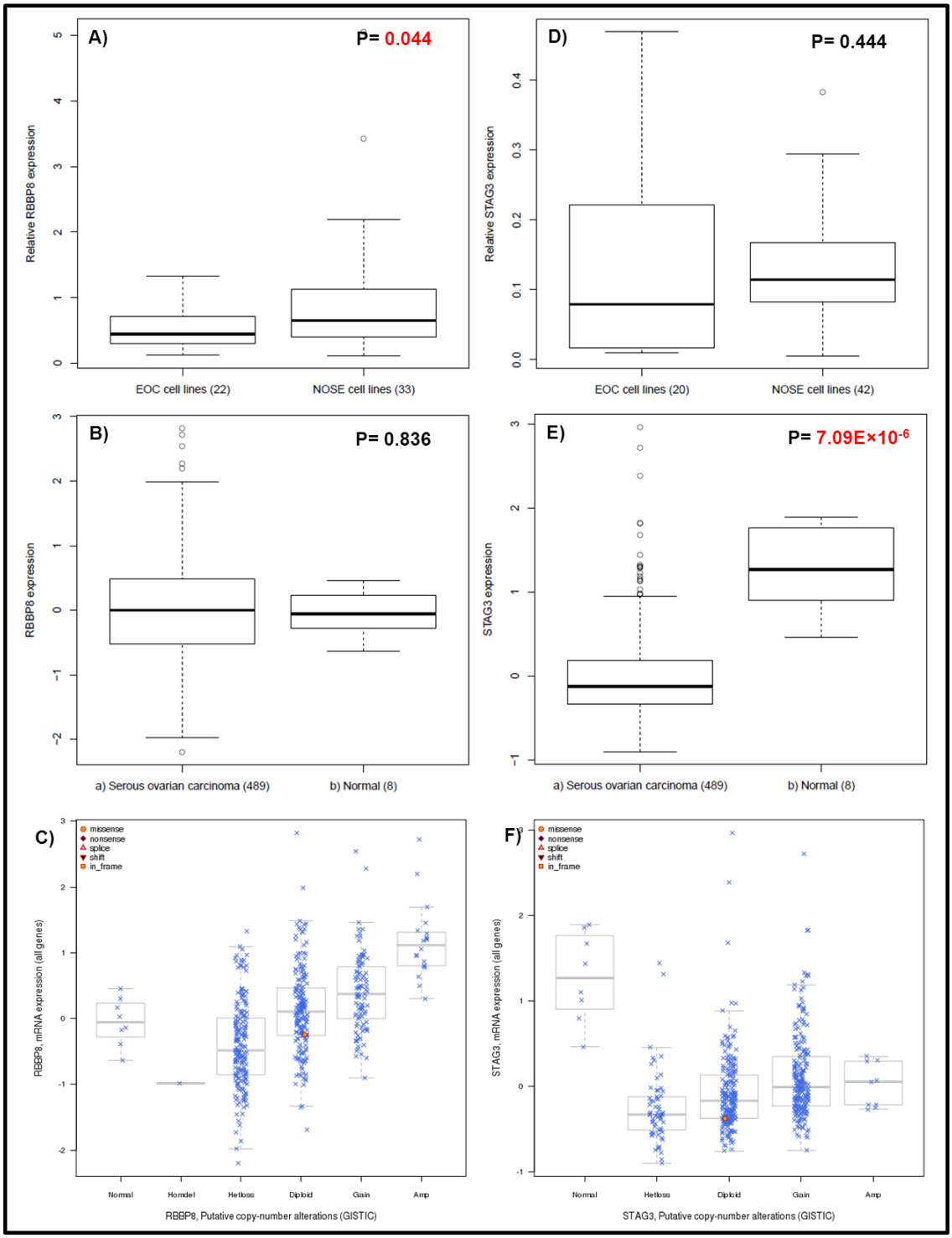


Figure 4.7: Differential expression of selected MMCT-18 genes between EOC and NOSE cell lines and between SOC and normal FT tissue. Differential expression between EOC and NOSE cell lines for: A) *RBBP8* , D) *STAG3*. Differential expression between SOC and normal FT for: B) *RBBP8*, E) *STAG3*. Differential expression relative to copy number variation between SOC and normal FT for: C) *RBBP8*, F) *STAG3*.

4.6 Discussion

Previous studies have demonstrated LOH in the chromosomal location 18q using microsatellite analysis (Cliby et al, 1993, Chenevix et al, 1992) and loss in chromosome 18q using CGH analysis (Arnold et al, 1996). Subsequent research has revealed tumour suppressor genes implicated with ovarian cancer development located in chromosome 18 and reported loss of their expression, including *SMAD4*, *SMAD2* and *DCC* and Cables (Dong et al, 2003). Previous CGH performed within our group has also reported loss of chromosome 18q in 55% of tumours analysed (Ramus et al, 2003) and microcell mediated chromosome transfer of chromosome 18 (MMCT-18) and subsequent microarray analysis had been performed in order to identify potential genes implicated in EOC development (Dafou et al, 2009). My project focused on nine candidate genes from several chromosomes that showed altered expression in the MMCT-18 project and aimed to identify any potential functional role of SNPs within these candidate genes in the EOC development and survival from ovarian cancer.

In the present study I evaluated LOH across genes *AIFM2, AKTIP, AXIN2, CASP5, FILIP1L, RBBP8, RGC32, RUVBL1* and *STAG3* in ovarian cancer. I have chosen to investigate LOH across the candidate genes using tSNPs across the genes to investigate whether genetic variants have a functional role in ovarian carcinogenesis by leading to LOH. Such functional role would be evidence for an alternative mechanism for selection of cells with growth advantage once they have lost one allele, rather than the traditional mechanisms such as mutations and epigenetic changes. To date numerous common genetic variants have been shown to be associated with the risks of several cancer types, but rarely has any functional rationale been established for their risk association. Therefore, I investigated whether tSNPs within all nine genes were somatically altered in primary ovarian cancers alongside an association study performed for the same tSNPs within these genes by Lydia Quaye within our group. Furthermore, I was interested to establish if there were any associations of LOH across the candidate genes with survival from ovarian cancer. I have also correlated the LOH frequencies I found with aCGH data

available within our group and I have examined the expression of all MMCT-18 candidate genes using expression data from TCGA for SOC and normal FT tissue or by assaying the differential expression of two selected genes in EOC versus NOSE cell lines. As previously discussed the results from these two expression models may show different roles for candidate genes in EOC development and is explained by differences both in the cancer and normal samples of the two models. Similarly that expression results may vary between the MMCT-18 microarray analysis results coming from only one of the hybrid cell lines used and the expression of a particular gene in EOC cell lines that represent several different histologies of ovarian cancer or SOC that are only high grade serous adenocarcinomas. Ideally, in the future, the number of EOC cell lines used in similar expression assays will be large enough to separate the analysis according to the subtypes of EOC the cell lines have originated from. A summary of the LOH and expression results and TCGA information for each gene is shown in Table 4.10.

The highest LOH frequency was found for gene *AXIN2,* and was 64%*.* This observation may suggest a loss of function role for *AXIN2* in ovarian cancer development. The mechanism for the high observed LOH for *AXIN2* did not rely on preferential loss of a specific allele of any of the genotyped SNPs (5 genotyped SNPs in tumour and matching germline DNA, with ratios of allele loss ranging between 1-1.5). However, it should be noted that preferential loss of an allele of another *AXIN2* SNP could be a possibility, because tSNPs do not provide LOH information for SNPs in LD with them but only genotype information.

AXIN2 is located in chromosome 17 and codes for a protein that controls axin formation during embryonic development. AXIN2 has been shown to be implicated in major pathways including the Wnt signalling, stress activated protein kinase (SAPK) and transforming growth factor β (TGFβ) pathways (Polakis, 2000, Salahshor et al, 2005). AXIN2 negatively regulates the Wnt signalling pathway by inducing the formation of a β-catenin destruction complex and is also involved in mitosis by influencing chromosome segregation. *AXIN2* aberrations including LOH, mutations and deletions have been identified in many forms of cancer including medulloblastoma, breast cancer, hepatocellular

carcinomas and oesophageal squamous cell carcinomas (Salahshor et al, 2005). The deregulation of β-catenin is a reported feature in ovarian endometrioid carcinomas and mutations in *AXIN2* have been shown to contribute to that, indicating a tumour suppressor role for *AXIN2* (Wu et al, 2001). In agreement with an *AXIN2* tumour suppressor role specifically in ECs is the fact that it was up-regulated only in hybrid TOV112D (endometrioid carcinoma cell line) of the MMCT-18 project and not in the TOV21G (clear cell line) hybrid. Interestingly, the univariate survival analysis I performed showed that the LOH found in *AXIN2* was strongly associated with worse survival of the non-SC tumour group containing ECs (HR=1.59 (1.09-2.30), P=0.001), which is a result that remained significant after Bonferroni correction. The fact that the association did not remain significant after prognostic factor adjustment is a result of the strong association of LOH in *AXIN2* with the stage on ovarian cancer and should be taken into account that adjustments in that case might be masking real association of LOH with survival. If anything, the fact that LOH in *AXIN2* is such an important indicative marker of EOC stage is an interesting finding of this study as loss of chromosome 17 has been previously associated with stage in EOC. Previous studies have shown that LOH in chromosome 17 in ECs may indicate transition to a more aggressive tumour (Shenson et al, 1995). Additionally, another study has reported that high frequencies of chromosome 17 loss were associated with high grade tumours and when stage was used as the outcome variable in multiple logistic regression analysis, chromosome loss was significantly associated with worse survival (Pieretti et al, 2002).

High LOH was also found in gene *AKTIP* with overall LOH of 46%. AKTIP, also known as FT1, is widely distributed in adult tissues and has been shown to induce apoptosis regulating the function of Protein Kinase B (PKB)/Akt activity (Remy et al, 2004). The high observed LOH was consistent with a 29% loss observed by aCGH in 68 of the ovarian tumours used for the LOH analysis in cytoband 16q12.2 where AKTIP gene was located suggesting a loss of function role for *AKTIP* in EOC development consistent with the up-regulation of the gene in the MMCT-18 hybrids. To my knowledge this is the first report showing high frequency if LOH in *AKTIP* in ovarian cancer and there is no further literature suggesting any AKTIP aberrations found in ovarian cancer.

Mutational analysis by Sanger Institute revealed no mutations of this gene in breast carcinomas (http://www.sanger.ac.uk/perl/ genetics/CGP/ cosmic) and only 0.2% of SOC studied revealed *AKTIP* somatic mutations according to TCGA.

Gene	Chromosomal locus	Epxression parental vs hybrid cell lines	Expression in SOC vs normal FT tissue (TCGA)	Expression in NOSE vs EOC cell lines	% LOH	Ascociation of LOH with survival (UV/MV)	aCGH	Mutations in SOC (TCGA)
AIFM2	10q22.1	3 fold up in TOV112D & TOV21G hybrids	Upregulated in SOC	N/A	27	All subtypes*/ None	27% gain	None
AKTIP	16q12.2	4 fold up in TOV112D & TOV21G hybrids	No difference	N/A	46	None	29% loss	None
AXIN2	17q23-q24	5 fold up in TOV112D hybrids	No difference	N/A	64	All subtypes,Non serous/ None	No change	0.2% Somatic
CASP5	11q22.2-q22.3	7 fold up in TOV21G hybrids	No difference	N/A	33	All subtypes*/ None	21% loss	0.6% Somatic
FILIP1L	3q12.1	5 fold up in TOV112D hybrids	No difference	N/A	32	None	No change	0.2% Somatic
RBBP8	18q11.2	7 fold up in TOV112D hybrids	No difference	Downregulated in EOC cell lines	23	All subtypes, Serous*/ all subtypes *	18% gain	0.2% Somatic
RGC32	13q14.11	3 fold down in TOV112D hybrids	No difference	N/A	34	None	29% loss	None
RUVBL1	3q21	25 fold down in TOV112D & TOV21G hybrids	No difference	N/A	30	None	31% gain	None
STAG3	7q22.1	9 fold up in TOV21G hybrids	Dowregulated in ovarian tumours	No difference	19	None	56% gain	0.2% Somatic

Table 4.10: Summary of results for the nine MMCT-18 candidate genes. UV: Univariate, MV: Multivariate, (*): did not remain significant after Bonferroni correction.

The observed allelic imbalance, recorded as LOH for *AIFM2* gene was 27%. *AIFM2* is a gene coding for a mitochondria associated protein implicated in caspase-independent apoptosis. Previous work has suggested that *AIFM2* has a promoter that is activated by p53 and has been found down-regulated in a majority of human tumours of several tissues including kidneys, colon, cervix and ovary (Wu et al, 2004). Interestingly, we found that the region containing *AIFM2* had a 28% copy number gain contradicting the observed LOH and up-regulation on the hybrids that indicate a tumour suppressor role for this gene. Also contradicting the loss of function role proposed by the LOH and up-regulation in the hybrids was the over-expression of *AIFM2* I found in SOC compared to normal FT tissue (P=0.009) after analysis of TCGA mRNA

expression data. It is interesting that the increased expression of *AIFM2* in SOC was attributed mainly to copy number gains found in the SOC samples that were over-expressed compared to normal according to TCGA. The different roles for the genes proposed by the different models studies can have various explanations. The gain detected by CGH in contrast with the loss detected by LOH analysis, may indicate that the LOH observed is due to amplification of the remaining allele rather than deletion of the lost one. SOC are high grade serous ovarian carcinomas and the hybrid cell lines are non-serous cell lines. Thus the difference in the expression of the gene in the two models may indicate that *AIFM2* may have a distinct role in EOC development for serous and another for non-serous subtypes. Another plausible explanation for the contradiction may be based on previous observations that show that some cancer cells have up-regulated expression of a mutated protein and dominant-negative protein as is the case for p53 for many tumours (Lutzker et al, 1996). It would be interesting to perform mutational analysis for *AIFM2* in the samples where gain was identified by aCGH to test such a hypothesis. However, there are no reported mutations in *AIFM2* in a panel of 489 SOC analysed by TCGA.

CASP5, an up-regulated gene in the TOV21G hybrids, was found to exhibit 33% LOH. The region of *CASP5* also demonstrated a 21% loss by aCGH further supporting its tumour suppressor role in ovarian cancer. *CASP5* is a member of the cysteine aspartic acid protease family and acts as an apoptosis regulator. I did not identify any AS LOH for the *CASP5* tSNPs I genotyped. *CASP5* has been reported to be occasionally mutated in human cancers that demonstrate microsatellite instability including gastric, colon, lung (Soung et al, 2008), leukaemia (Takeuchi et al, 2003) and endometrial cancer (Schwarz et al, 1999). According to TCGA only 0.6% of the SOC studied were found to have a somatic mutation. *CASP5* expression is down-regulated in non-small cell lung cancer and has been proposed to be a tumour suppressor gene that when inactivated contributes in lung carcinogenesis and promotes the highly metastatic phenotype (Gemma et al, 2001). I found that LOH in *CASP5* was associated with worse survival for ovarian cancer (HR=1.47 (1.04-2.08), P=0.031) in the univariate analysis. Even though the association did not remain significant after multiple testing correction. The fact that no association was

found in none of the subgroups when the subtypes were stratified in serous and non-serous samples indicates that as the sample size of the study increases the association becomes apparent. Additionally, the association did not remain after adjusting for prognostic factors. The association of *CASP5* LOH with survival from EOC, if strengthened by larger sample size studies, could be explained with CASP5 being implicated in chemotherapy response. CASP5 has recently been reported to be implicated with the outcome of combination chemotherapy in a study that showed that by using a PI3-Kinase inhibitor in combination with Adriamycin, a leukaemia chemotherapeutic drug, the levels of CASP5 significantly increased leading to a major increase in the chemosensitivity of cancer cells (Suvasini et al, 2010).

FILIP1L with 32% LOH identified, was shown to be up-regulated in the hybrids, thus a candidate tumour suppressor gene for ovarian cancer. The region corresponding to the gene did not demonstrate significant copy number changes by aCGH, similarly to the genes *AXIN2* and *RBBP8,* which although they exhibited high frequency of LOH they did not show similar changes by aCGH. The reason why frequent LOH in these candidate genes did not correlate with the aCGH may be because of the different sensitivities of the two methods for detecting allelic loss. More probable is that LOH analysis may have detected allelic deletions that were accompanied by the conversion or duplication of the remaining allele, resulting in allelic imbalance without a change in copy number.

FILIP1L is coding for filamin A interacting protein and is involved in cell proliferation and migration. I have not identified any allele specific loss in any of the tSNPs genotyped within this gene, thus other mechanisms may be responsible for the observed LOH. Recent research reports decreased expression of *FILIP1L* in ovarian cancer cell lines compared to normal ovarian epithelial cell lines. This study provides evidence supporting that the observed down-regulation is due to hypermethylation of a promoter which is related to the invasive phenotype in ovarian cancer (Burton et al, 2011). The therapeutic potential of this gene was pursued by a group that recently proposed a gene therapy treatment for ovarian cancer after using polymeric nanoparticles to deliver functional FILIP1L in ovarian cancer cell lines *in vitro* and then to *in vivo*

models and successfully achieving an effective growth inhibition of the tumour and reversed neoplastic phenotype (Xie et al, 2011).

RBBP8 is coding for a retinoblastoma binding protein which regulates cell proliferation. RBBP8 has been shown to control double strand break resection leading to homologous recombination (Huertas et al, 2009). Its expression is regulated post transcriptionally possibly by differential efficiency of its RNA translation at different stages of the cell cycle. RBBP8 expression varies with the cell cycle progression in a manner identical to BRCA1, peaking at the G1/S checkpoint control. RBBP8 interacts with BRCA1, CTBP1 (C-terminal binding protein 1) and RB1 and it is conceivable that it serves as a regulatory link between these three distinct pathways of tumour suppression (Yu et al, 2000, Chinnadurai, 2006). The association of *RBBP8* with ovarian cancer risk is inconclusive. Whereas one recent study has shown association of *RBBP8* haplotypes and breast cancer risk in a cohort of 2,825 *BRCA1* mutation carriers (Rebbeck et al, 2011), no such association with ovarian cancer risk in *BRCA1/2* mutation carriers was reported (Rebbeck et al, 2009). We have found associations of tSNPs in *RBBP8* with risk in developing EOC in a cohort of sporadic ovarian tumours when the analysis was restricted to serous ovarian cancer cases in a study of 829 SCs and 2895 controls [HetOR = 0.83 (0.70–0.98), HomOR = 0.80 (0.63–1.03), *p*-trend = 0.032] (Notaridou et al, 2011). I found 23% overall LOH in *RBBP8* in accordance with MMCT-18 hybrid up-regulation indicating *RBBP8* has a tumour suppressor role in ovarian cancer. In the univariate survival analysis I found an association of *RBBP8* LOH with worse survival of serous ovarian cancer patients [HR=2.09 (1.23-3.50), P=0.005] although the result did not remain significant after Bonferroni correction. The association of *RBBP8* LOH with worse survival of all EOC subtypes patients was significant after Bonferroni correction (HR=2.48 (1.60-3.84), P<0.00001). Both recorded associations were not significant after Bonferroni in the multivariate analysis after adjusting for clinical factors. Similar to *AXIN2*, it is conceivable that the multivariate negative results for *RBBP8* revealing no such association could be positive results masked by the fact that stage and LOH in those genes are so closely associated. In parallel, an association study by Lydia Quaye has reported two tSNPs in *RBBP8* being

associated with ovarian cancer survival (Quaye et al, 2009). Taking into account these observations and the biological function of RBBP8 it is intriguing to hypothesise that the association of *RBBP8's* LOH with worse survival of ovarian cancer patients might be a true observation and could be related to absence of the gene product conferring resistance to chemotherapy. This hypothesis is strengthened by research that has reported significant down-regulation of RBBP8 in tamoxifen resistant breast cancer cells and further showed that silencing *RBBP8* was associated with acquired tamoxifen resistance in breast cancer (Wu et al, 2007). I have evaluated the expression of *RBBP8* in EOC and NOSE cell lines and found that it was under-expressed in EOC cell lines. Expression data from the TCGA revealed no difference in the expression of *RBBP8* between ovarian tumour and normal FT tissue. This indicates that *RBBP8* under-expression in EOC cell lines may be implicated in chemoresistance for particular subtypes of EOC and depending on the cell type they originate from. Additionally, lower expression of *RBBP8* by stromal cells within the FT tissue used for the TCGA expression data may mask a possible higher expression in FT epithelial cells. Future work to evaluate the expression differences between chemotherapy resistant and sensitive EOC cell lines of different subtypes could elucidate *RBBP8's* possible role in chemosensitivity for ovarian cancer patients.

RUVBL1 was down-regulated in the hybrids and aCGH showed a 31% gain at the region where the gene is mapped. Interestingly, LOH analysis showed a frequency of 30% LOH for *RUVBL1*. However, as LOH could just be allelic imbalance the observed LOH might be due to amplification of the remaining allele. We have also shown that two SNPs in *RUVBL1*, rs13063604 and rs7650365, were associated with increased risk of serous ovarian cancer in a parallel association study performed by Lydia Quaye (Notaridou et al, 2011). RUVBL1 is involved in chromatin remodelling. It is essential for transcriptional regulation of proto-oncogenes and cell proliferation and plays an essential role in oncogenic transformation by *MYC* (Wood et al, 2000). This biological feature fits tightly with the observation of *RUVBL1* being down-regulated in the hybrids indicating a gain of function role in ovarian cancer development. However, the precise functions of this protein are not fully understood and it is possible that in

different tissues or even different subtypes of tumour cells it might acquire distinct functions. For example, recent work has shown that RUVBL1 is modulating the foci formation by RAD51 following DNA damage by double-strand DNA breaks and crosslinks (Gospodiniv et al, 2009). Thus, it is conceivable that RUVBL1 may be influencing the response of tumour cells to DNA crosslinking drugs such as cisplatin. I am unable to draw a more substantial hypothesis on the role of *RUVBL1* in EOC using the data from this project and information in the literature. More research is needed to elucidate the oncogenic molecular pathways RUVBL1 is implicated in to fully understand its role in EOC development.

I also made conflicting observations for the gene *RGC32* in cytoband 13q14.11 for which aCGH data showed 30% loss and LOH analysis showed 34% LOH, but it was down-regulated in TOV112D hybrids. RGC32 is involved in cell cycle regulation and it is induced by p53 in response to DNA damage but its role in carcinogenesis is still controversial. RGC32 has been reported to be over co-expressed with mitotic marker Ki-67 in colon, breast and prostate cancer indicating that is required for growth in tumour cells by deregulating the cell cycle (Fosbring et al, 2005). A recent study has reported that *RGC32* was down-regulated via promoter methylation in non-small lung cancer and the methylation status of the tumours was associated with worse survival in patients with functional p53 and favourable survival in cases with *TP53* mutations (Kim et al, 2010). The latter finding suggests that it may be interesting to conduct a study in the future that will look into the LOH frequency and expression of *RGC32* in ovarian tumours and EOC cell lines respectively stratifying them according to their *TP53* mutational status. *RGC32* was not found to be differentially expressed between SOC and normal FT according to the mRNA data from TCGA or between EOC and NOSE cell lines.

STAG3, on chromosome 7q22.1, encodes a component of the meiosis specific cohesion complex (Pezzi et al, 2000, Prieto et al, 2001). STAG3 is not expressed in embryonic stem cells that form follicle like ovarian structures and is thought to contribute to the inability of those cells to progress through meiosis (Ivana Novak et al, 2006). STAG3 has also been shown to be activated in lymphoma cells with *TP53* mutations induced by irradiation (Martins Kalejs et al,

2006). Other studies have implicated mutations within *STAG3* with chromosomal instability observed in colorectal cancers (Thomas et al, 2007), and the gene is associated with chromosome segregation and down regulation in testicular cancer (Skotheim et al, 2005). *STAG3* was up-regulated in TOV21G hybrids. In agreement to this, I found that there was 19% LOH in *STAG3* in ovarian tumours. Interestingly, I did not find any differential expression of *STAG3* between NOSE and EOC cell lines. The latter result did not correlate with the up-regulation found in the hybrids or with the expression data from TCGA showing higher expression of *STAG3* in normal FT tissue compared to SOC. Whether this difference was due to expression of stromal cells in the FT tissue driving the observed increased expression in normal compared to SOC or was a very strong indication that STAG3 may be implicated only in the development of high grade serous EOCs originating from FT cells remains to be investigated. However, I am inclined to propose the latter. A subtype specific role of *STAG3* in EOC development was also suggested by the increased expression of *STAG3* found only for the hybrid cell line TOV21G and not in TOV112D, however both of the cell lines are non-serous cell lines which in turn is puzzling.

The most interesting finding regarding *STAG3* was the identification of a tSNP within the gene that was responsible for allele specific LOH. SNP rs1637001 in the *STAG3* gene showed significant non-random allelic specific imbalance by LOH analysis in tumours (94% loss of allele A, P=0.015). rs1637001 is a SNP located at the 3'UTR of *STAG3* gene. This SNP tags 15 other SNPs within *STAG3* gene, 15 of those being in intronic regions or within the 3' UTR, and one located in exon 24. However as mentioned before, LOH information is given only about the specific SNP, and the fact that this tags others does not gives any information about their LOH status, but just genotype information.

The recorded allele specific LOH for this SNP was found by aCGH to be the likely result of amplification of the minor G allele; amplification rather than deletion of an allele has previously been suggested as a potential mechanism to explain detected LOH at a SNP (LaFramboise et al, 2005). Supporting this hypothesis was the observation of *STAG3* down-regulation in SOC compared to

normal being characterised by a variety of copy number variations the majority of them gains (Figure 4.8 F).

The same SNP in this gene also showed evidence of association with disease risk in serous ovarian cancer cases in the parallel association study conducted by Lydia Quaye, although this was not confirmed in the imputed data from a very large GWAS later performed by our group and collaborators. The potential synergy between a putative risk association for a germline genetic variant and the preferential somatic amplification of one of the alleles during tumour development is intriguing and further investigation would be worthwhile. There are several examples in the published literature of allele specific imbalance of polymorphic markers in primary tumours. For example, in one study, analysis of the *DAL1* gene in breast cancer found that 94% of tumours showing LOH retained the C allele of the C2166T SNP; in another study, 73% of lung tumours with LOH involving the *P34* gene retained the G allele of the A106G SNP; a third study reported 83% of breast tumours with loss of the Pro allele of Arg72Pro in the *P53* gene. (Mao et al, 2006, Kittiniyom et al, 2004, Wang et al, 2007, Wegman et al, 2009). However, synergy between genetic risk alleles and somatic alterations has not been reported before.

What is unclear for the rs1637001 variant located in the 3'UTR of *STAG3*, is whether or not the preferential allelic amplification targets rs1637001 or *STAG3* specifically. The amplified region detected by aCGH in tumours extended across several genes including *STAG3*; it is not clear that the amplification is functionally relevant to ovarian cancer development though. More detailed *in vitro* cell biology studies would be interesting to set up in the future to address this. The biological significance of the preferential allele loss could also be further evaluated by performing genotype specific gene expression analysis using RNA from NOSE and FTE cell line bank that I have established. Another way to assess functionality of alleles would be by performing protein-DNA binding studies after transfecting with different constructs expressing each allele.

Conclusion

In conclusion, in this study I have investigated and found frequent LOH for tSNPs across the candidate genes selected from the MMCT-18 project. I have identified functional evidence suggesting allele specific imbalance for somatic genetic alterations in primary ovarian tumours, for the common allele of a SNP in the *STAG3* gene. I also found associations of LOH across *AXIN2*, *RBBP8*, *CASP5* and *AIFM2* with survival of EOC patients but remained significant only for *RBBP8* after adjusting for prognostic factors and none remained significant after Bonferroni correction. Finally, this study suggests that LOH in the candidate MMCT-18 genes *AXIN2*, *BRCA1*, *RGC32*, *RBBP8* and *AKTIP*, is a strong indicative marker of EOC stage and/or grade.

♣ Publications containing work from this chapter

Notaridou M*, Quaye L*, Dafou D, Jones C, Song H et al. “Common alleles in candidate susceptibility genes associated with risk and development of epithelial ovarian cancer”. 2011. Int J Cancer. 128(9):2063-74

Quaye L*, Dafou D*, Ramus SJ, Song H, Gentry-Maharaj A, **Notaridou M**, Hoqdall E et al. “Functional complementation studies identify candidate genes and common genetic variants associated with ovarian cancer survival”. 2009. Hum Mol Genet.18(10):1869-78.

* Authors contributed equally to this work

5 Functional analysis of candidate genes and SNPs identified from an ovarian cancer Genome Wide Association Study

5.1 Introduction

Over the last few years, genome wide association studies (GWAS) have discovered more than 200 new common low-penetrance susceptibility loci for several types of cancer. The variants identified confer low risks for disease and thus the functional role of the low-risk loci identified in the development of the disease is presumably modest. It is a significant challenge to understand how these loci contribute to the development of cancer at the biological level and to unravel the links between the identified variants and the molecular basis of risk etiology. This area of research is part of post-GWAS characterisation of risk loci and focuses on understanding the way in which low-risk variants lead to the initiation of cancer pathogenesis and will be beneficial for the more effective screening, prevention and treatment of the disease.

An ovarian cancer GWAS performed previously within our lab and collaborators has identified 6 susceptibility regions significantly associated with susceptibility to ovarian cancer ($P< 10^{-6}$) (Song et al, 2009, Goode et al, 2010, Bolton et al, 2010). The risk variants identified are mapped to loci to chromosome loci 2q31, 3q25, 8q24, 9p22, 17q21 and 19p13 (Table 5.1). Most of the identified SNPs are non-coding thus do not affect the sequence of any target gene's translated protein. My hypothesis is that the risk variants at these loci may be acting as *cis*-regulatory variants. *Cis*-regulatory variants are SNPs either within a gene or up to 1Mb proximal to the start or up to 1Mb distal to the end of the gene. *Trans*-regulatory variants are SNPs that are located elsewhere in the genome. I hypothesised that the identified SNPs in the GWAS may have a functional role in ovarian cancer development by regulating in *cis* the transcriptional output of candidate genes located up to a 1Mb region around the most significant SNP or the methylation status of the candidate genes. I expanded this hypothesis by investigating whether genes mapping in the risk

loci could be implicated in EOC development using the NOSE & FTE versus EOC model established and described in chapter 3.

Locus	Associated SNP	SNP Location & Function	Stages 1, 2 and 3 combined				Gene(s) in closest proximity to associated SNP
			All subtypes		Serous cases only		
			OR (95% CI)	P-value	OR (95% CI)	P-value	
2q31**	rs2072590	Within non-coding gene, unknown	1.16 (1.12-1.21)	4.5×10^{-14}	1.2 (1.14-1.25)	3.8×10^{-14}	*HOXD1* *HOXD3*
3q25 **	rs2665390	Gene intron, unknown function	1.19 (1.11-1.27)	3.2×10^{-7}	1.24 (1.15-1.34)	7.1×10^{-8}	*TIPARP*
8q24**	rs10088218	Intergenic non coding, unknown function	0.84 (0.80-0.89	3.2×10^{-9}	0.76 (0.70-0.81)	8×10^{-15}	*MYC*
9p22*	rs3814113	Intergenic non coding , unknown function	0.82 (0.79-0.86)	5.1×10^{-19}	0.77 (0.73-0.81)	4.1×10^{-21}	*BNC2*
17q21**	rs9303542	Gene intron, unknown function	1.11 (1.06-1.16)	1.4×10^{-6}	1.14 (1.09-1.20)	1.4×10^{-7}	*SKAP1*
19p13***	rs8170	Coding, non synonymous	1.12 (1.07-1.17)	2×10^{-6}	1.2 (1.13-1.27)	6×10^{-10}	*MERIT40*
	rs2363956	Coding, non synonymous	0.96 (0.89-0.94)	4×10^{-7}	0.96 (0.89-0.94	7×10^{-11}	*ANKRD41*

Table 5.1: SNPs found in association with epithelial ovarian cancer susceptibility and genes in closest proximity. OR (95% CI): per allele odds ratio (95% confidence interval). The P values presented for all the samples and serous only samples are from the combined 3 stages of the GWAS *(Song et al, 2009), ** (Goode et al, 2010), *** (Bolton et al, 2010).

Aims and objectives:

1. A pilot study to test the hypothesis that candidate genes within the associated loci may have a role in EOC will be initially performed. For this study, the protein coding gene in the closest proximity to each of the identified most significant variants will be selected. A follow up study using a larger panel of normal (combined NOSE & FTE) and EOC cell lines will then be performed to validate the results obtained from the pilot study. This will also be an extended study in the sense

that the differential expression of more genes within the risk loci (located up to 1Mb of associated SNPs) will be assayed.

2. The functional role of the significantly associated with EOC susceptibility SNPs in the transcriptional regulation of the candidate genes will be investigated. Initially, I will evaluate gene expression in NOSE cell lines for the candidate genes in the closest proximity with the associated SNPs relatively to their genotype. In an extended study I will investigate genotype specific gene expression using a larger number of normal (NOSE & FTE) cell lines for all the additional candidate genes selected within 1Mb distance from the associated SNPs.
3. The potential *cis*-regulatory function of the EOC associated SNPs will also be investigated by evaluating whether there are associations between SNP genotype and methylation status of the candidate genes. Initially a pilot analysis will be performed using methylation and genotype data from the blood of healthy individuals. The methylation status of CpG islands of the closest to the associated SNPs candidate genes will be compared between the different genotype groups. A follow up study will then be performed using a larger number of samples to validate any genotype specific methylation identified and check whether the methylation status of the additional candidate genes selected within 1Mb distance from the SNPs is associated with their genotype.

5.2 Evaluating differential expression of candidate genes between NOSE and EOC cell lines - Pilot study

The most significant low-risk SNPs identified by the GWAS were in the closest proximity with the protein coding genes *HOXD1* and *HOXD3* for locus 2q31, *TIPARP* for locus 3q25, *MYC* for locus 8q24, *BNC2* for locus 9p22, *SKAP1* for locus 17q31 and *MERIT40* (or *BABAM1*) and *ANKRD41* (or *ANKLE1)* for locus 19p13 (Figure 5.1). *BABAM1* and *ANKLE1* are the more recently assigned official names for the last genes but the alias names *MEIRT40* and *ANKRD41* will be used in this thesis.

The known biological function of some of the candidate genes made the investigation for their possible role in EOC attractive. *HOXD1* and *HOXD3* are members of the homeobox D family genes that are transcription factors regulating genes involved in morphogenesis, differentiation and development. *TIPARP* (TCDD-inducible poly (ADP-ribose) polymerase) is a member of the PARP polymerase superfamily and may play role in adaptive response to chemical exposure. *MYC* is a well-established oncogene that activates the transcription of growth related genes. *BNC2* (basonuclin 2) is a transcription factor that is specific for skin keratinocytes and may play a role in differentiation of spermatozoa and oocytes. *SKAP1* (src kinase associated phosphoprotein 1) is positively regulating T-cell receptor signalling by enhancing the MAP kinase pathway. *MERIT40* (Mediator of RAP80 interactions and targeting subunit of 40 kDa) has the most relevant interesting reported function as it is a component of the BRCA1 interacting complex involved in homologous recombination at the sites of DNA double strand breaks (DSBs). *ANKRD41* (ankyrin repeat domain 41) is not well characterised but may be involved at DNA cleavage during DNA damage response (http://www.genecards.org).

The differential expression of the candidate genes in closest proximity to the risk variants was assessed between 48 NOSE and 24 EOC cell lines. The cell lines were cultured as described in the methods and RNA was extracted. The RNA was then reversed transcribed into cDNA and the cDNA amplified using Taqman Real time expression assays in the ABI7900 Taqman Sequence detection system.

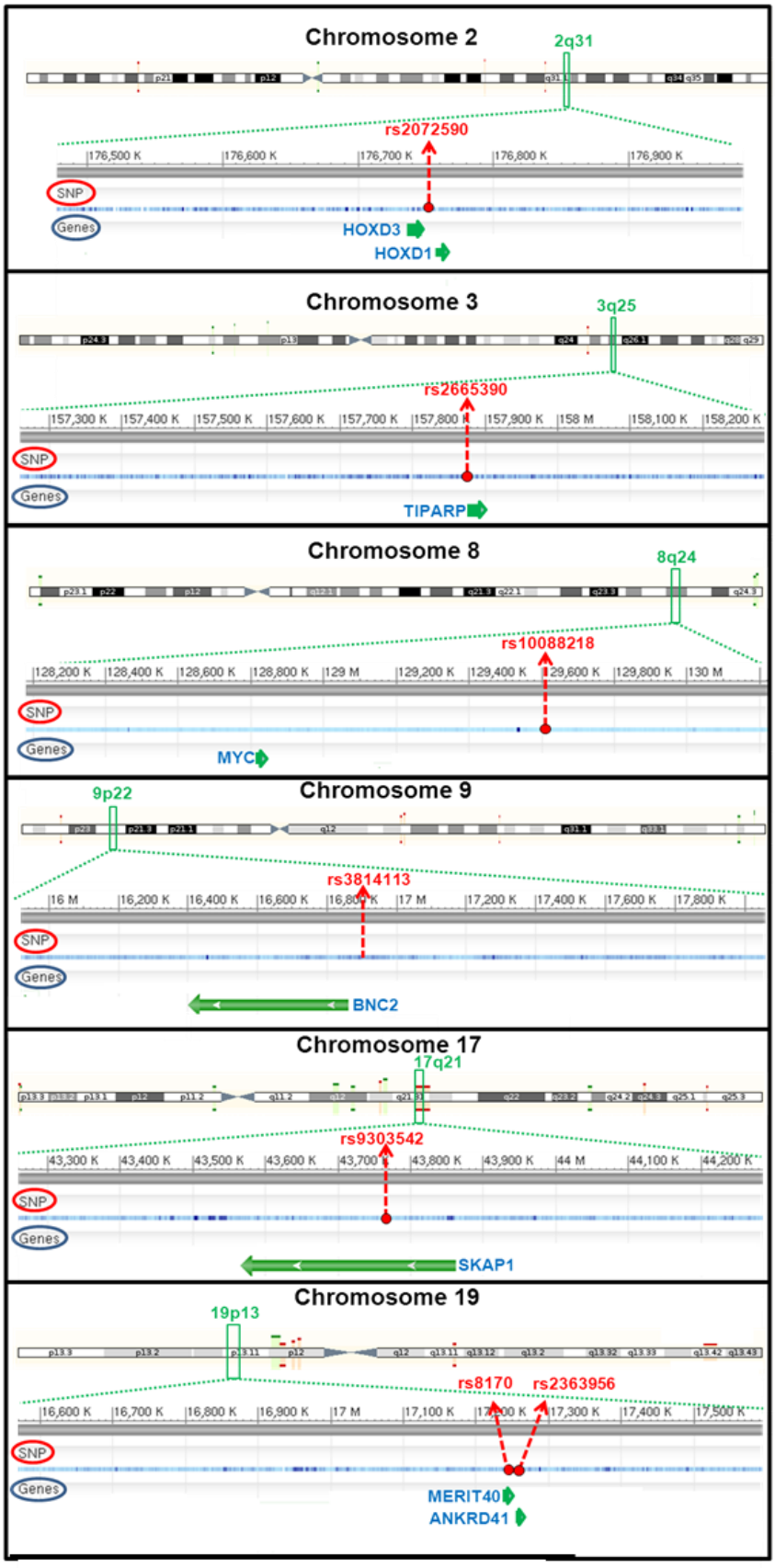


Figure 5.1: Mapping of the low-risk susceptibility SNPs identified by the GWAS at chromosomes 2, 3, 8, 9, 17 and 19. The genes in closest proximity to those SNPs within loci 2q31, 3q25, 8q24, 9p22, 17q21 and 19p13 selected to be investigated are illustrated.

5.2.1 Quality control and candidate gene expression analysis

To confirm the linear relationship of expression versus RNA quantity for each target probe and endogenous control probe a standard curve was created for each assay using serial dilutions of a mix of the cDNAs in the study 5x more concentrated than the samples to be run. Data were determined to be reliable if an excellent fit (correlation coefficient R^2 between 0.95 and 1) of the standard curve data to a straight line was observed or at least being able to create a suboptimal line (R^2>0.8). The standard curves for each of the probes for the candidate genes and the 2 endogenous genes are shown at Figure 5.2.

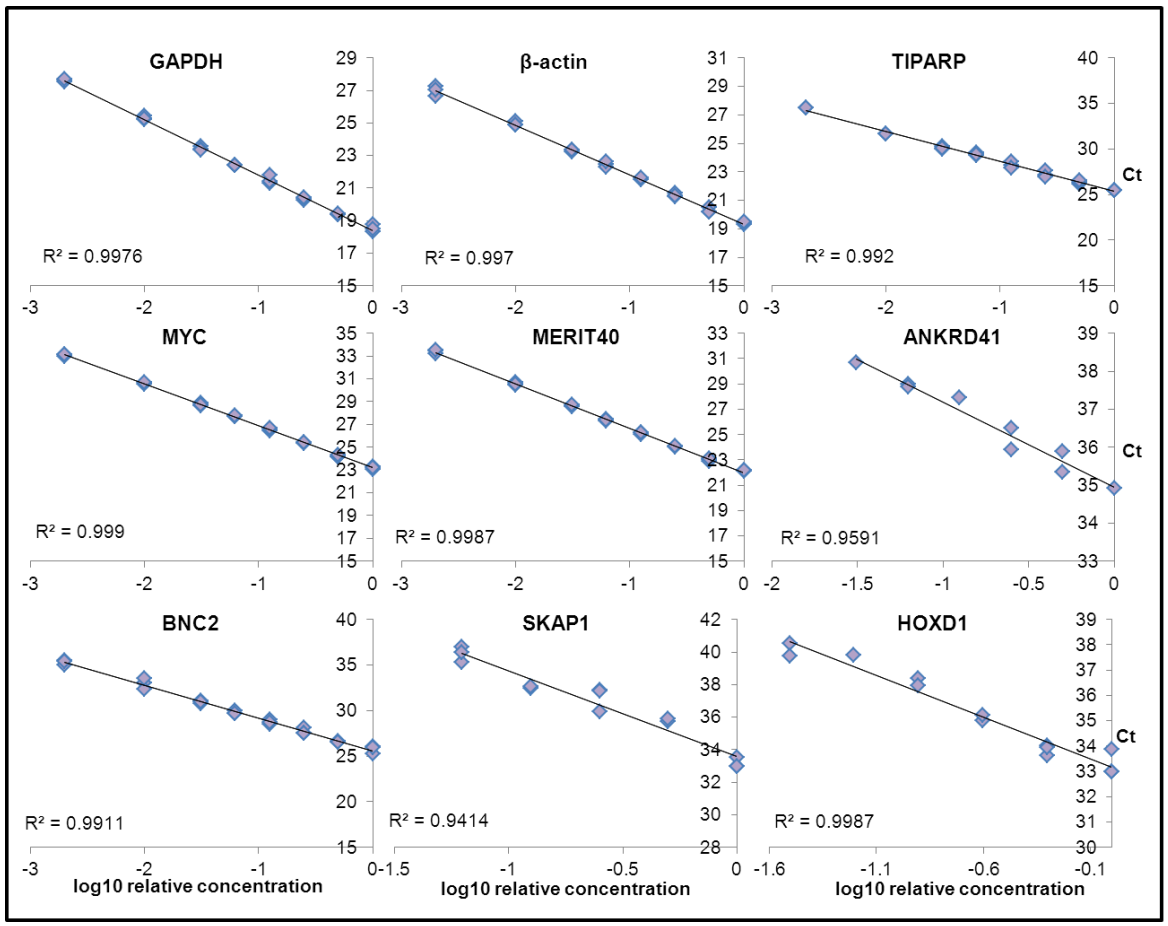


Figure 5.2: Standard curves of the candidate gene and endogenous control probes- Pilot study

To test the reliability of the assay, I initially looked at the relative expression of GAPDH normalised to β-actin to exclude the possibility of the results being biased by the relative expression of the endogenous controls. The expression of the endogenous controls was not different between NOSE and

EOC cell lines as shown at Figure 5.3. The assays were performed in two sets and alongside each set the endogenous controls were also run.

The Taqman Real time expression data were analysed using the comparative ΔΔCt method and the expression values of all cell lines were generated relative to either the lowest or highest expression of a NOSE cell line normalized against GAPDH and β-actin. For samples where the standard deviation was >0.6 between the replicates, the replicate causing the deviation was removed. Differences in the relative expression of each candidate gene between EOC cell lines and NOSE cell lines were assessed using the nonparametric two sided Wilcoxon Rank sum test using R software and P values were generated. Values were considered statistically significant if P<0.05.

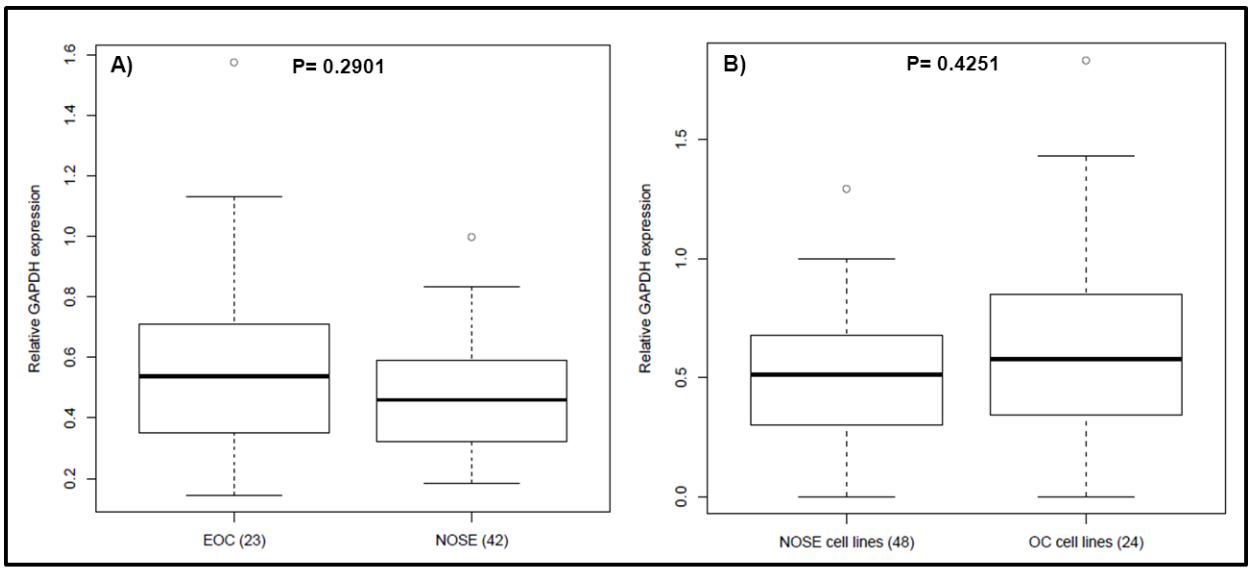


Figure 5.3**:** GAPDH expression in NOSE versus EOC cell lines normalised to β-actin. Used as endogenous controls for the assays for a) *BNC2*, *MERIT40*, *TIPARP*, *MYC* genes b) *HOXD1*, *SKAP1*, *ANKRD41* genes

5.2.2 Results of Real time PCR expression analysis of candidate genes

The expression data from NOSE versus EOC cell lines were compared with publicly available microarray data from the TCGA between serous ovarian carcinomas (SOC) and normal fallopian tube tissue (FT) which were analysed as previously described in Chapter 3. Additionally, within our group a novel EOC transformation model was recently created. Briefly, an immortalised NOSE cell line (IOSE) was transformed by introducing *MYC* and *MYC/KRAS*

(Lawrenson et al, 2011). Thus, I have also compared the differential expression between NOSE versus EOC cell lines for the candidate genes with the differential expression between IOSE and IOSE-*MYC* & IOSE-*MYC/KRAS* cell lines from data generated in a microarray analysis performed in the context of that project.

Locus 2q31: Candidate genes *HOXD3* and *HOXD1*

For the 2q31 locus, I tested the expression of both the *HOXD3* and *HOXD1* genes. Both were in the closest proximity with the risk SNP at this locus. However, *HOXD3* expression in both EOC and NOSE cell lines was too low for reliable detection. *HOXD1* expression was significantly higher in EOC cell lines compared to NOSE cell lines as shown by the boxplots in Figure 5.4 A&B ($P<10^{-8}$). This observation suggests that *HOXD1* has a gain of function role in EOC development. This observation was not confirmed in the IOSE transformation model. There was no differential expression of *HOXD1* between the IOSE and IOSE_*MYC* and ISOS-*MYC/KRAS* cell lines. The observed gain of function role of *HOXD1* in ovarian cancer development was also not supported by TCGA data where its expression levels did not differ between SOC and normal FT (Figure 5.4 C and D respectively).

Locus 3q25: Candidate gene *TIPARP*

The *TIPARP* gene which contains in one of its introns the most significant SNP of locus 3q25, exhibited significantly reduced expression in EOC compared to NOSE cell lines for both endogenous controls as shown in Figure 5.5 A&B ($P<10^{-8}$). The observed lack of expression for *TIPARP* in EOC cell lines suggests a loss of function role for *TIPARP* in ovarian cancer development. A loss of function role for *TIPARP* in EOC is also suggested with the higher *TIPARP* expression observed in the IOSE cell line compared to its transformed counterparts (Figure 5.5 C).This finding was also supported by a significantly reduced expression of the gene in SOC compared to normal FT after I performed the analysis of mRNA data available from TCGA ($P= 1.11\times10^{-4}$, Figure 5.5 D).

Locus 8q24: Candidate gene *MYC*

SNP rs10088218, the most significant SNP in locus 8q24 is not located in very close proximity with any genes but is found in a gene desert. *MYC* gene, which is ~760kb away of rs10088218, was not the gene in closest proximity to this SNP but was indeed one of the two genes more closely located to the SNP. *PVT1* was the other gene, which is a non-protein coding gene. The reason I selected to initially investigate *MYC* is because it is a known oncogene with a well-established role in the development of EOC and additionally it is a protein coding gene. I found that *MYC* was significantly over-expressed in EOC compared to NOSE cell lines and this observation was consistent with the well-established role of the gene as an oncogene. (Figure 5.6 A&B). *MYC* was used for the transformation of the IOSE in the EOC transformation model but when co-transfected with *KRAS* its expression decreased to the initial levels of the control cell line (Figure 5.6 C) Interestingly, TCGA data showed that *MYC* expression levels didn't change between SOC and normal FT (Figure 5.6 D).

Locus 9p22: Candidate gene *BNC2*

SNP rs3814113 in locus 9p22 was the most significantly associated SNP with EOC risk in the GWAS ($P= 5.1\times10^{-19}$). The gene in the closest proximity to this SNP was *BNC2* (~30kb away). After evaluating the expression in EOC cell lines compared to NOSE cell lines, a loss of function role of *BNC2* in EOC is supported by the findings. The gene was found to be significantly under-expressed in EOC cell lines compared to NOSE ($P= \sim1\times10^{-5}$, Figure 5.7 A&B). This is an interesting result supported by the observation of being significantly under-expressed in the transformed IOSE cell lines compared to the control ($P<0.01$, Figure 5.7 C). I also found that the expression of *BNC2* was significantly reduced in SOC compared to normal FT in expression data from the TCGA ($P= 1.6\times10^{-4}$).

Locus 17q21: Candidate gene *SKAP1*

For chromosome 17, locus 17q21, the most significant SNP rs9303542 is located in a *SKAP1* intron. I found that there was significantly increased expression of *SKAP1* in EOC cell lines compared to NOSE cell lines as shown

at Figure 5.8 A&B with a P value between 10^{-4} and 10^{-6} depending on the endogenous control used for normalisation. This trend is indicating a gain of function role for *SKAP1* in EOC development. Interestingly, a significant differential expression was observed between normal FT and SOC but with a different trend suggesting a loss of function role in EOC since the expression of *SKAP1* appears to be reduced in SOC compared to normal FT (Figure 5.8 D).

Locus 19p13: Candidate genes *MERIT40* and *ANKRD41*

The most significant SNPs of locus 19p13 were rs8170 and rs2363956 which are located in the coding regions of genes *MERIT40* and *ANKRD41* respectively. *MERIT40* expression was significantly higher in EOC cell lines compared to NOSE cell lines as shown at Figure 5.9 A&B ($P<10^{-5}$). These data suggest that *MERIT40* has a gain of function role in ovarian cancer development. This finding was supported by the increased expression I found in serous ovarian adenocarcinoma compared to normal FT. *MERIT40* did not appear differentially expressed in the IOSE transformation model (Figure 5.9 C). *ANKRD41* expression was not significantly different between EOC and NOSE (Figure 5.10 A&B).

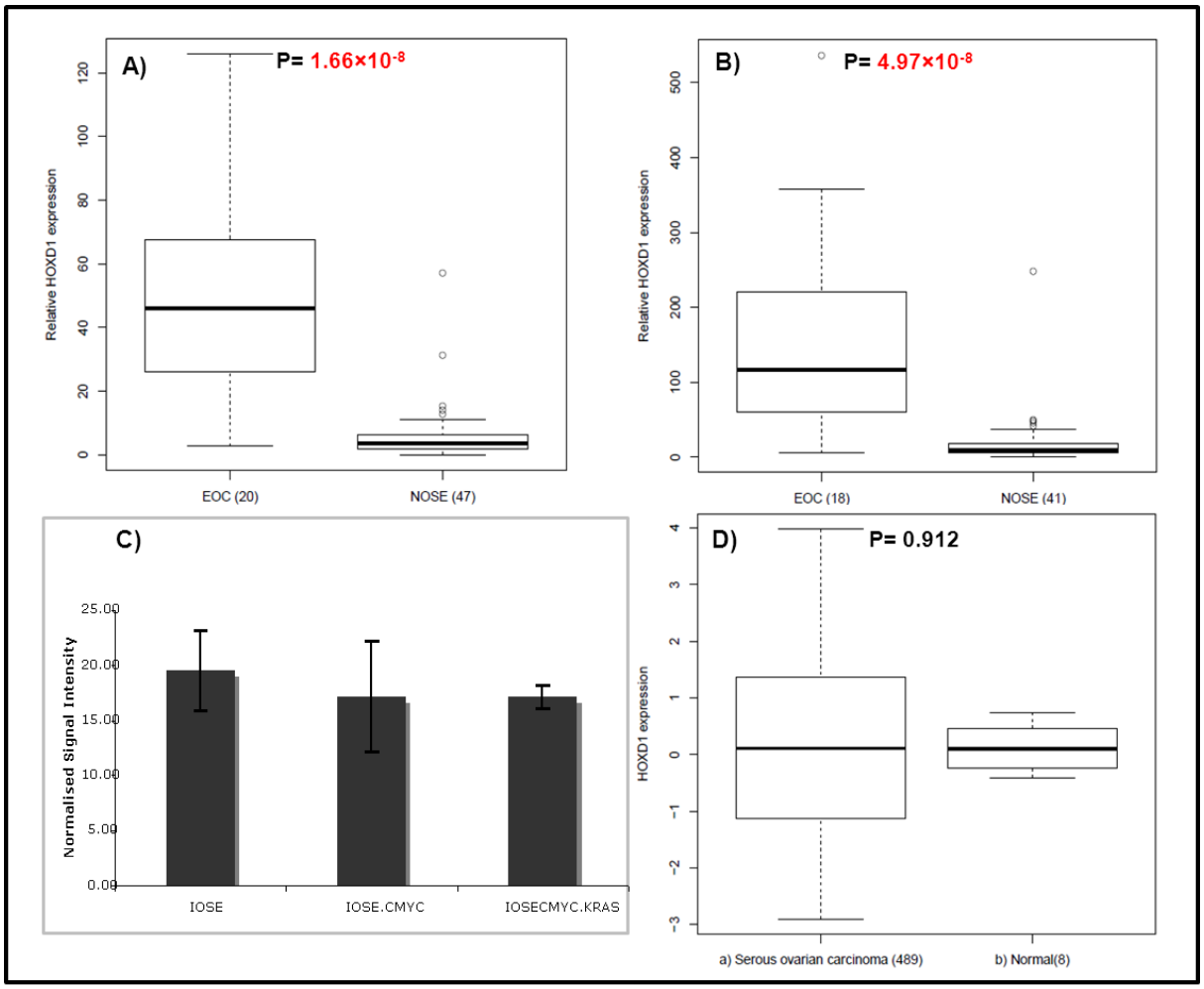


Figure 5.4**:** *HOXD1* expression in three differential expression models. *HOXD1* expression in EOC versus NOSE cell lines normalised to A) β-actin and B) GAPDH. C) *HOXD1* expression in the IOSE transformation model. D) *HOXD1* expression in SOC versus normal FT.

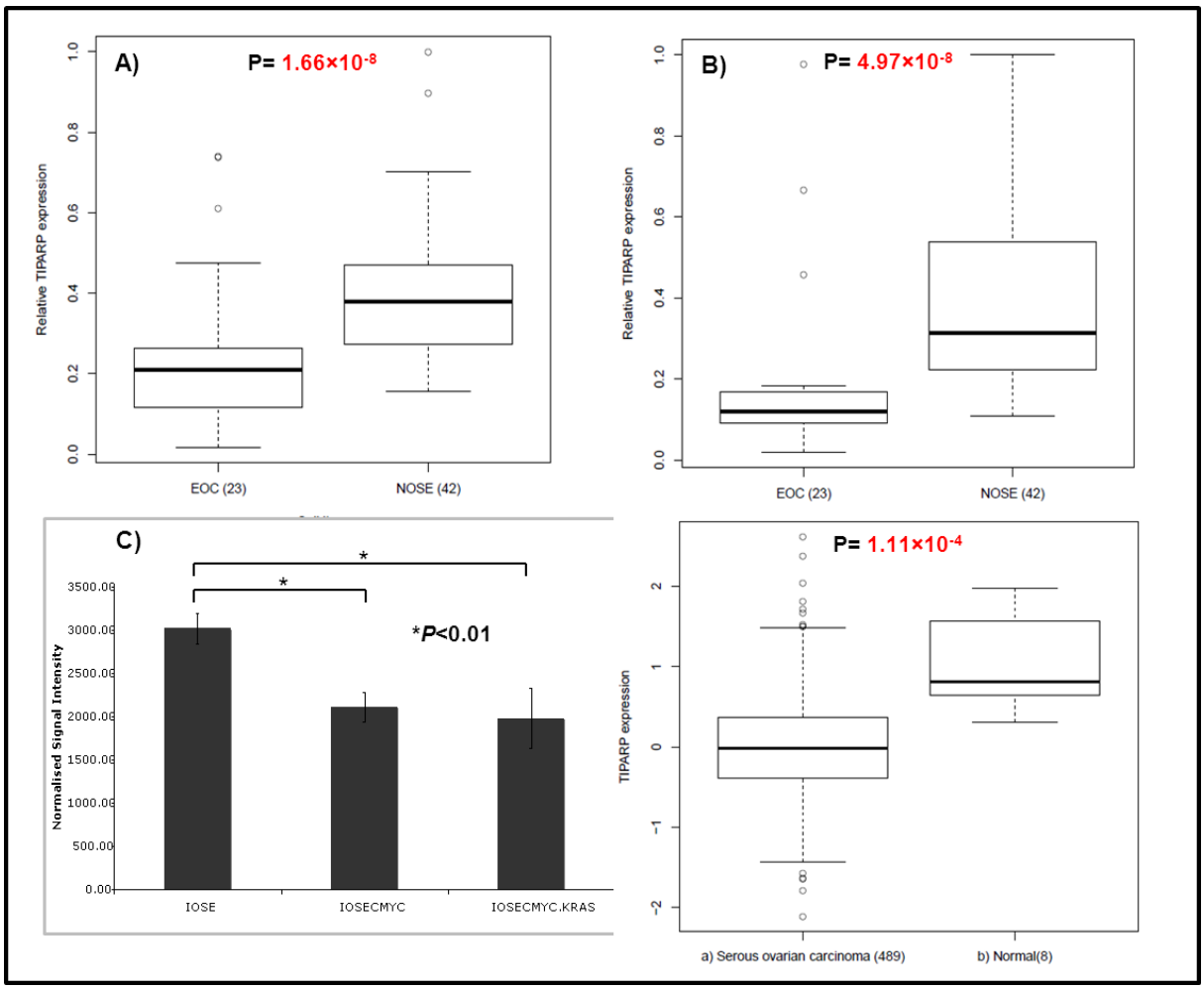


Figure 5.5: *TIPARP* expression in three differential expression models. *TIPARP* expression in NOSE versus EOC cell lines normalised to A) β-actin and B) GAPDH. C) *TIPARP* expression in the IOSE transformation model. D) *TIPARP* expression in SOC versus normal FT.

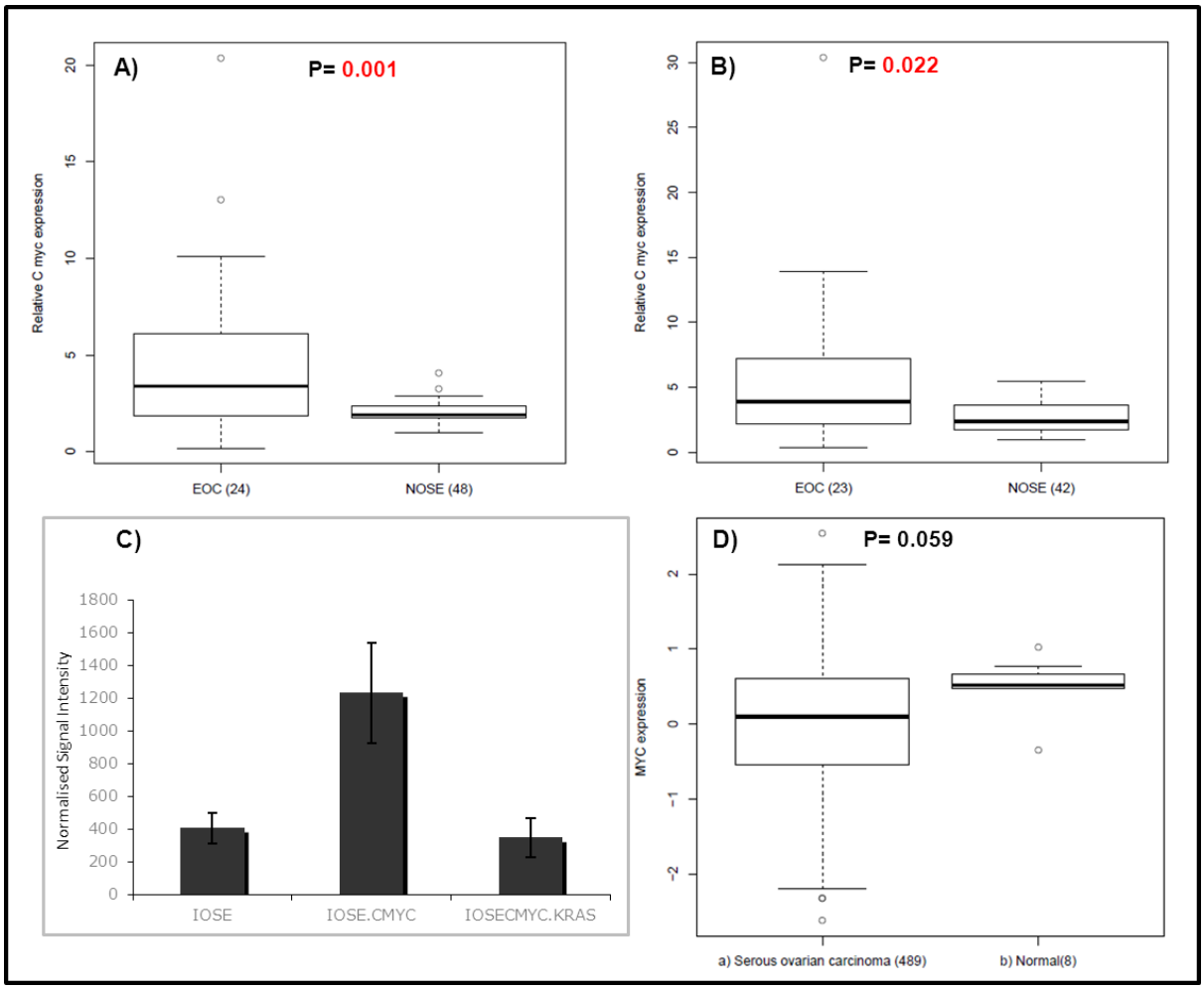


Figure 5.6: *MYC* expression in three differential expression models. *MYC* expression in NOSE versus EOC cell lines normalised to A) β-actin and B) GAPDH. C) *MYC* expression in the IOSE transformation model. D) *MYC* expression in SOC versus normal FT.

.

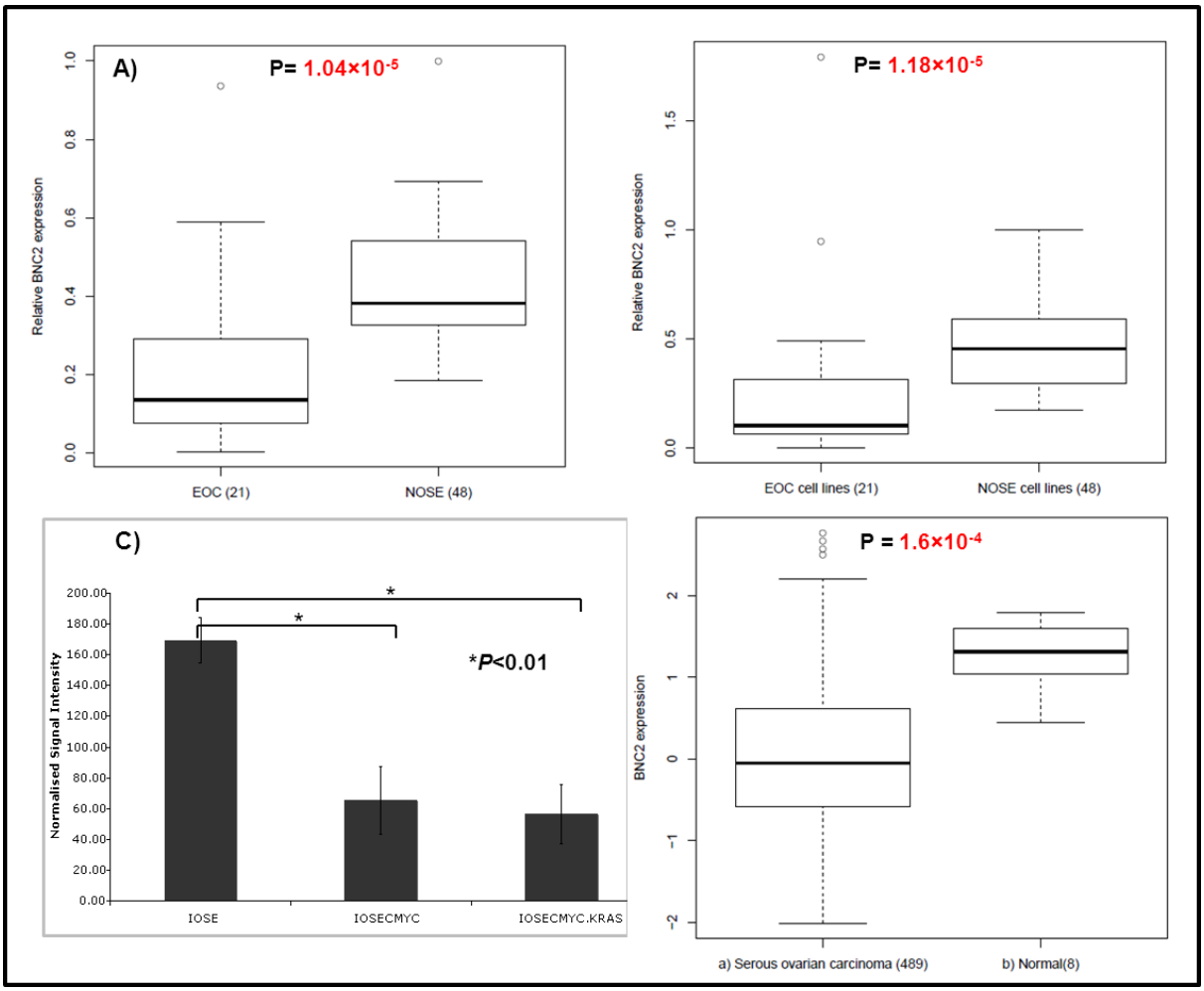


Figure 5.7**:** *BNC2* expression in three differential expression models. *BNC2* expression in NOSE versus EOC cell lines normalised to A) β-actin and B) GAPDH. C) *BNC2* expression in the IOSE transformation model. D) *BNC2* expression in SOC versus normal FT.

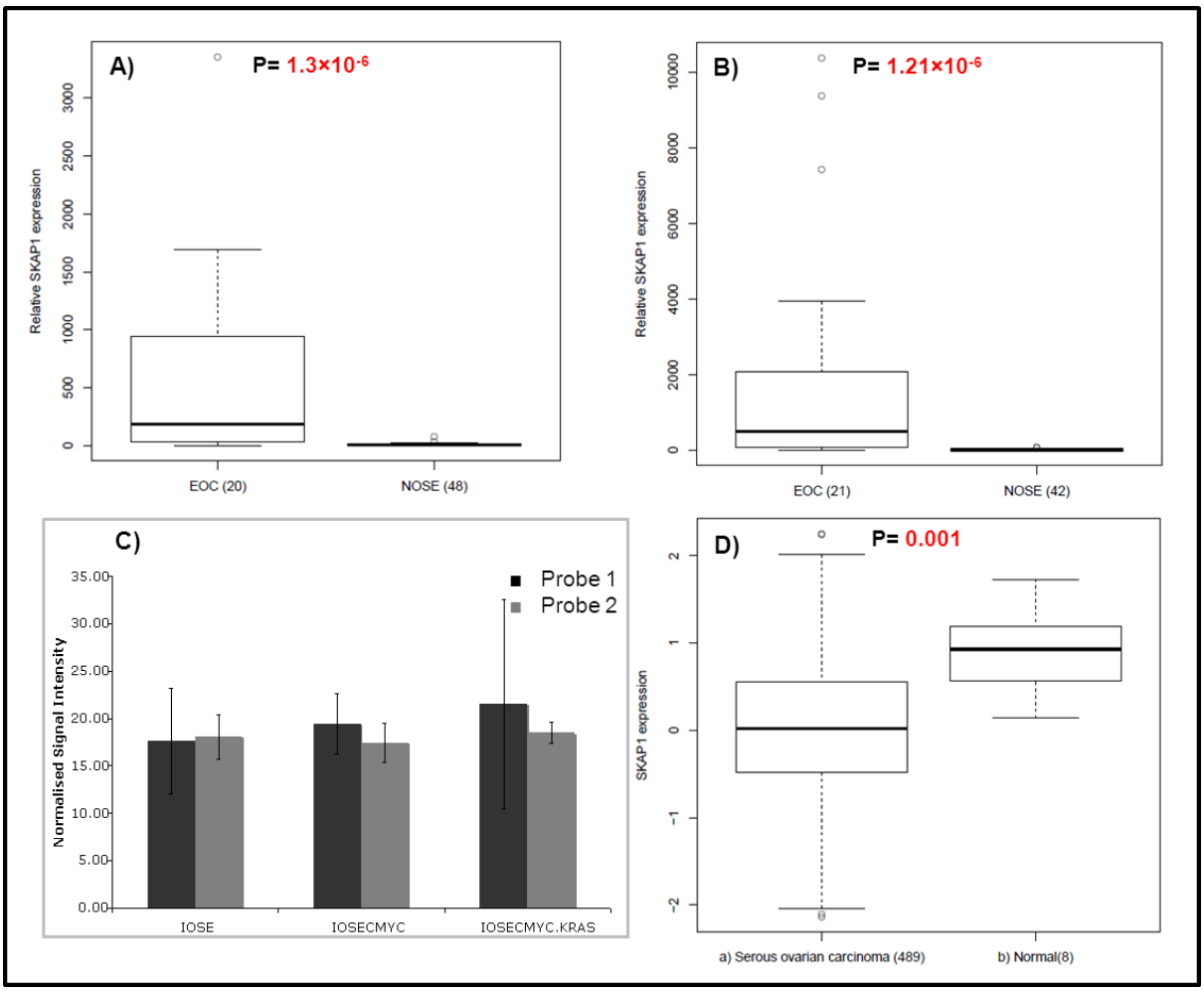


Figure 5.8**:** *SKAP1* expression in three differential expression models. *SKAP1* expression in NOSE versus EOC cell lines normalised to A) β-actin and B) GAPDH. C) *SKAP1* expression in the IOSE transformation model. D) *SKAP1* expression in SOC versus normal FT.

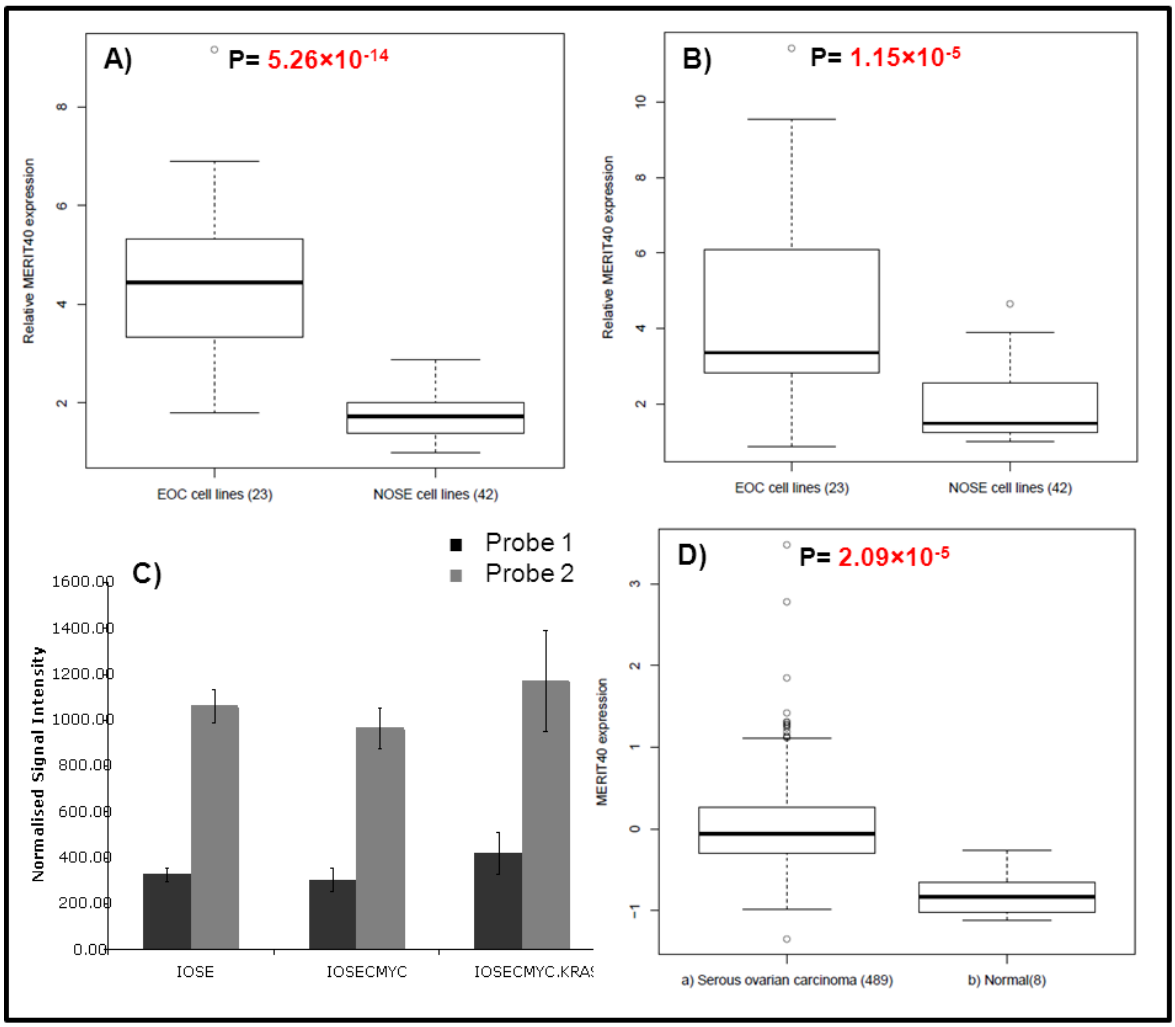


Figure 5.9: *MERIT40* expression in three differential expression models. *MERIT40* expression in NOSE versus EOC cell lines normalised to A) β-actin and B) GAPDH. C) *MERIT40* expression in the IOSE transformation model. D) *MERIT40* expression in SOC versus normal FT.

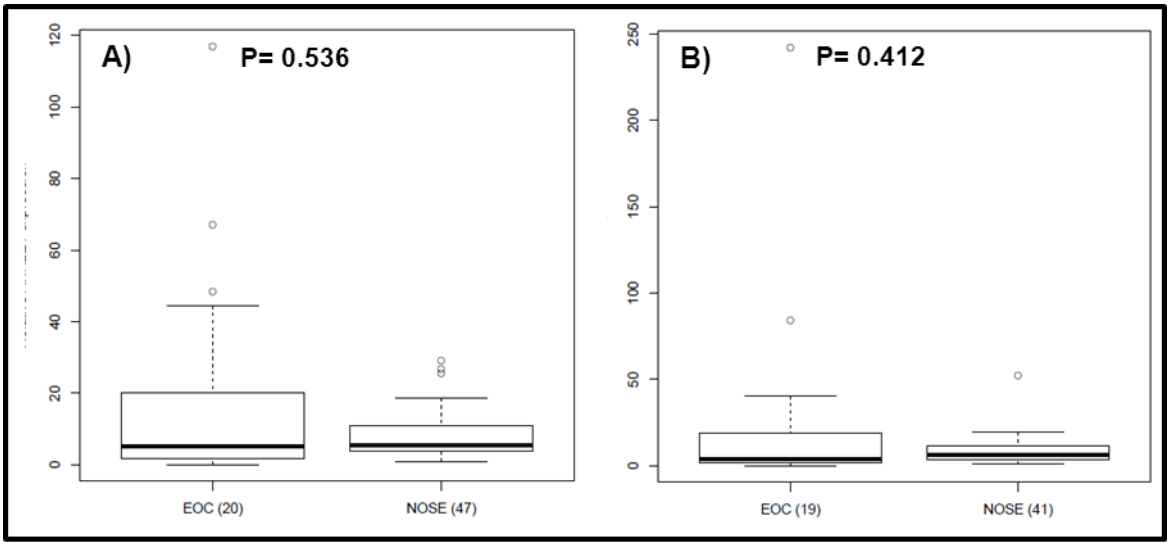


Figure 5.10: *ANKRD41* expression in NOSE versus EOC cell lines. *ANKRD41* expression normalised to A) β-actin and B) GAPDH,

5.3 Evaluating differential expression of an extended list of candidate genes between normal (NOSE & FTE) and EOC cell lines- Extended study

The initial approach of investigating the potential role of candidate genes emerging from the GWAS in the development of ovarian cancer was to examine the genes that were in the closest proximity to the most significant SNPs at the six loci. However, it cannot be ruled out that there are other genes at these susceptibility regions that may be functionally important in EOC. Therefore, I expanded these studies to include a larger panel of normal (combined NOSE & FTE) and EOC cell lines and evaluated additional genes in each region. A map for each of the chromosomal loci of interest was established using Genome Built 36.3 in Ensemble (http://www.ensembl.org/index.html) and more genes around the significant SNPs were selected located in less than 1Mb distance away of either side of the SNPs (Table 5.2).

Locus	Associated SNP	Alleles (Major/ Minor)	Position of SNP (Genome Build 36.3) (bp)	Size of region for selected genes (Kb)	Candidate genes selected
2q31**	rs2072590	G/**T**	176750879	1025	12
3q25 **	rs2665390	T/**C**	157880443	1600	6
8q24**	rs10088218	G/**A**	129613131	1150	2
9p22*	rs3814113	T/**C**	16905021	515	2
17q21**	rs9303542	A/**G**	43766499	1025	23
19p13***	rs2363956	G/**T**	17255124	980	15

Table 5.2**:** Regions established for selection of candidate genes around the most significant EOC risk variants. The number of additional candidate genes selected to investigate their differential expression between normal and EOC cell lines is also tabulated.

5.3.1 Fluidigm study design

Sixty genes were selected to be investigated in total, 12 at chromosome 2q31, 6 at chromosome 3q25, 2 at chromosome 8q24, 2 at chromosome 9p22, 23 at chromosome 17q21, and 15 at chromosome 19p13. A brief description of the function of those genes is shown in Table 5.3. Differential mRNA expression of the selected genes was evaluated using the cDNA of normal cell lines comprised by 59 NOSE and 5 FTE cell lines compared to 45 EOC cell lines. As previously described in chapter 3, NOSE217 had not been included in the analysis because it was found to express really high levels of FSP in the immunofluorescence characterisation, an indication of stromal contamination. FT05 cell line was also removed from the analysis as it was found to have an abnormal karyotype. The assays used were Taqman gene expression assays by Applied Biosystems run as a multiplex using the 96.96 dynamic array chips by Fluidigm. The protocol and assay selection is described in more detail in the Methods section of this thesis. The endogenous controls run were β-actin and GAPDH as well as TBP. Each sample was run in triplicate. A pooled sample of the cell lines used was created to generate standard curves for each of the assay also run in triplicate. The 60 assays were separated run in 3 different chips (sets) each including the endogenous control genes for the expression of the candidate genes from the relevant chip to be normalised against.

5.3.2 Quality control analysis

Extensive quality control (QC) analysis was performed in order to ensure that the data were reliable. Firstly, the pass rate for each of the assays was calculated using the 64 normal and 45 EOC cell lines (total n= 109). Assays where the pass rate was less than 80% were considered as failed. However, even if the accuracy of statistical analysis is important, the biological nature of the experiment should not be overlooked. Therefore, for assays where the pass rate was less than 80% the cell type specific pass rate was calculated separately for normal and EOC cell lines. This is justified because a gene may be found differentially expressed between normal and EOC cell lines to such an extent that there is no detectable expression in one of the cell types. If that is the case it will be picked up as a failed assay in the overall quality control

analysis. Two assays that had an overall <80% pass rate, for genes *HOXD9* and *HOXB9*, were found to have a cell type specific expression and these genes were not considered as they failed assay pass rate since they demonstrated a pass rate of >80% in at least one of the cell types. 36 assays remained after removing assays with <80% pass rate. This was not considered to be alarming because very low overall pass rates could indicate that a gene is not normally expressed neither in normal (NOSE and FTE) nor ovarian cancer cell lines and not necessarily mean that the assay failed.

A standard curve was generated for each of the assays to ensure the linear relationship of sample amount and relative gene expression. Assays for which the standard curve had an R^2 <0.8 were failed. The standard curves for all assays are presented in Appendix 3, Figure 1. Two more assays were excluded from the analysis at this step of QC. The standard curves for all assays are shown in Appendix 3, Figure 1.

The reproducibility between the replicates is extremely important for testing the reliability of the gene expression data. The standard deviation of the Ct values was assessed between the triplicates for each assay. For each sample, when the standard deviation was >0.6, the replicate causing the deviation was removed as long as two replicates remained to be analysed. Samples where only one replicate had worked were failed. For some samples where one of the replicates had failed a standard deviation of <1 was accepted. This less stringent cut off for the standard deviation of <1 was used in order to keep two replicates in for the analysis and assays that the rate of this occurring was more than 15% were failed. 32 assays remained after this final step of assay quality control analysis. The quality control analysis for all genes chosen to evaluate their differential expression in the six loci is shown in Table 5.4.

Quality control analysis was also performed for the samples. Samples that had a pass rate of <85% were removed. 2 NOSE cell lines and 1 EOC cell line were removed. The remaining number of samples after QC was 57 NOSE, 5 FTE cell lines and 44 EOC cell lines. The endogenous controls assays β-actin and GAPDH passed for genes that were run in sets/chips 3 and 4. TBP failed for all chips and β-actin failed for set 2.

Chromosome	Gene	Locus	Function	Fluidigm Chip
2	KIAA1715	2q31	May be involved in limb and CNS development	Set 3
	EVX2	2q31.1	TF for the Homeobox genes	Set 2
	HOXD13	2q31.1	TF involved in morphogenesis, differentiation and development	Set 3
	HOXD12	2q31.1	TF involved in morphogenesis, differentiation and development	Set 3
	HOXD11	2q31.1	TF involved in morphogenesis, differentiation and development	Set 3
	HOXD10	2q31.1	TF involved in morphogenesis, differentiation and development	Set 3
	HOXD9	2q31.1	TF involved in morphogenesis, differentiation and development	Set 3
	HOXD8	2q31.1	TF involved in morphogenesis, differentiation and development	Set 3
	HOXD4	2q31.1	TF involved in morphogenesis, differentiation and development	Set 3
	HOXD3	2q31.1	TF involved in morphogenesis, differentiation and development	Set 3
	HOXD1	2q31.1	TF involved in morphogenesis, differentiation and development	Set 4
	MTX2	2q31.1	Transporter of proteins into the mitochondrion	Set 3
3	KCNAB1	3q26.1	Accessory potassium channel protein	Set 3
	SSR3	3q25.31	Associated with protein translocation across the ER membrane.	Set 4
	TIPARP	3q25.31	May play a role in adaptive response to chemical exposure	Set 4
	LOC730091	3q25.31	Hypothetical protein Unknown function	Set 3
	PA2G4P4	3q25.31	Proliferation associated pseudogene with Unknown function	Set 3
	LEKR1	3q25.31	Unknown function	Set 3
8	MYC	8q24.21	Activates the transcription of growth-related genes	Set 3
	PVT1	8q24	Potential oncogene	Set 4
9	BNC2	9p22.32	TF specific for skin keratinocytes with a role in the differentiation	Set 2
	CNTLN	9p22	Unknown function	Set 2
17	SP6	17q21.32	Promotes cell proliferation	Set 4
	SP2	17q21.32	Binds to GC box promoters elements and selectively activates mRNA synthesis	Set 4
	PNPO	17q21.32	Catalyzes the oxidation of either PNP or PMP into PLP	Set 4
	ATAD4	17q21.32	Transmembrane protein of unknown function	Set 2
	CDK5RAP3	17q21.32	Potential regulator of CDK5 activity. May be involved in cell proliferation.	Set 2
	COPZ2	17q21.32	Cytosolic protein associated with protein transport from the ER	Set 2
	NFE2L1	17q21.3	Activates erythroid-specific, globin gene expression	Set 3
	CBX1	17q21.32	Component of heterochromatin. Involved in epigenetic repression.	Set 2
	SNX11	17q21.32	May be involved in several stages of intracellular trafficking	Set 4
	SKAP1	17q21.32	Positively regulates T-cell receptor signalling by enhancing the MAP kinase pathway	Set 4
	HOXB1	17q21.3	Sequence-specific TF for the developmental regulatory system	Set 2
	HOXB2	17q21-q22	Sequence-specific TF for the developmental regulatory system	Set 3
	HOXB3	17q21.3	Sequence-specific TF for the developmental regulatory system	Set 3
	HOXB4	17q21-q22	Sequence-specific TF for the developmental regulatory system	Set 3
	HOXB5	17q21.3	Sequence-specific TF for the developmental regulatory system	Set 3
	HOXB6	17q21.3	Sequence-specific TF for the developmental regulatory system	Set 3
	LOC404266	17q21.32	Hypothetical protein Unknown function	Set 3
	HOXB7	17q21.3	Sequence-specific TF for the developmental regulatory system	Set 3
	HOXB8	17q21.3	Member of the Antp homeobox family TFs involved in development	Set 3
	HOXB9	17q21.3	Sequence-specific TF for the developmental regulatory system	Set 3
	PRAC	17q21	May play a regulatory role in the nucleus.	Set 4
	HOXB13	17q21.2	Sequence-specific TF for the developmental regulatory system	Set 2
	TTLL6	17q21.32	Glutamylase which preferentially modifies alpha-tubulin	Set 4
19	SIN3B	19p13.11	Transcriptional repressor. Anotagonizes MYC oncogenic activities	Set 4
	CPAMD8	19p13.11	Belongs to the complement component-3 involved innate immunity and DDR	Set 2
	HAUS8	19p13.11	Contributes to mitotic spindle assembly	Set 4
	MYO9B	19p13.1	May be involved in the remodelling of the actin cytoskeleton.	Set 3
	USE1	19p13.11	May be involved in targeting and fusion of Golgi-derived transport vesicles with the ER	Set 4
	OCEL1	19p13.11	Unknown function	Set 3
	NR2F6	19p13.1	Transcriptional repressor mainly involved in modulation of hormonal responses	Set 3
	USHBP1	19p13	Unknown function	Set 4
	MERIT40	19p13.11	Component of the BRCA1-A complex. May be involved in DNA damage repair at DSBs	Set 2
	ANKRD41	19p13.11	May induce DNA cleavage and DNA damage response	Set 4
	ABHD8	19p13.11	May have catalytic and hydrolase activity involved in metabolic processes	Set 2
	DDA1	19p13.11	May be involved in ubiquitination and proteasomal degradation of target proteins.	Set 2
	MRPL34	19p13.1	Assists in protein synthesis within the mitochondrion	Set 3
	TMEM16H	19p13.11	May act as a calcium-activated chloride channel	Set 4
	GTPBP3	19p13.11	May play a role in mitochondrial tRNA modification.	Set 2

Table 5.3**:** Selected candidate genes from the six EOC risk associated loci. Genes are shown in the order found across the 6 candidate loci (direction p→q arm). Highlighted in blue are the initially selected genes in the closest proximity to the most significant SNPs from the GWAS.

Chromosome	Gene	Assay pass rate (%) (109 samples)	% of samples per assay with SD>0.6	R^2 Standard Curve	Assay failed or passed
2	KIAA1715	99	0	0.978	PASSED
	EVX2	0	0	FAILED	FAILED
	HOXD13	22	5	0.868	FAILED
	HOXD12	5	2	0.921	FAILED
	HOXD11	8	3	FAILED	FAILED
	HOXD10	31	10	FAILED	FAILED
	HOXD9	73	17	0.876	FAILED
	HOXD8	84	7	0.878	PASSED
	HOXD4	45	13	0.667	FAILED
	HOXD3	61	10	0.828	FAILED
	HOXD1	28	6	0.936	FAILED
	MTX2	99	2	0.805	PASSED
3	KCNAB1	42	5	FAILED	FAILED
	SSR3	98	0	0.993	PASSED
	TIPARP	98	1	0.936	PASSED
	LOC730091	48	11	0.052	FAILED
	PA2G4P4	97	0	0.956	PASSED
	LEKR1	43	9	0.016	FAILED
8	MYC	100	0	0.953	PASSED
	PVT1	95	12	0.834	PASSED
9	BNC2	91	1	0.980	PASSED
	CNTLN	99	0	0.944	PASSED
17	SP6	32	11	0.189	FAILED
	SP2	98	0	0.991	PASSED
	PNPO	98	0	0.955	PASSED
	ATAD4	14	5	0.991	FAILED
	CDK5RAP3	97	5	0.978	PASSED
	COPZ2	86	3	0.953	PASSED
	NFE2L1	100	0	0.953	PASSED
	CBX1	100	1	0.977	PASSED
	SNX11	97	1	0.963	PASSED
	SKAP1	23	3	0.559	FAILED
	HOXB1	1	1	FAILED	FAILED
	HOXB2	78	15	0.205	FAILED
	HOXB3	82	10	0.954	PASSED

	HOXB4	87	6	0.809	PASSED
	HOXB5	73	13	0.943	FAILED
	HOXB6	84	8	0.925	PASSED
	LOC404266	15	2	0.108	FAILED
	HOXB7	88	4	0.989	PASSED
	HOXB8	30	7	0.050	FAILED
	HOXB9	79	5	0.943	PASSED (*)
	PRAC	6	4	FAILED	FAILED
	HOXB13	30	3	0.735	FAILED
	TTLL6	14	5	FAILED	FAILED
19	SIN3B	98	0	0.990	PASSED
	CPAMD8	14	3	0.653	FAILED
	HAUS8	97	3	0.982	PASSED
	MYO9B	100	0	0.983	PASSED
	USE1	97	1	0.906	PASSED
	OCEL1	95	8	0.890	PASSED
	NR2F6	98	1	0.944	PASSED
	USHBP1	7	5	FAILED	FAILED
	MERIT40	100	0	0.960	PASSED
	ANKRD41	68	18	0.717	FAILED
	ABHD8	99	2	0.922	PASSED
	DDA1	100	0	0.938	PASSED
	MRPL34	99	2	0.972	PASSED
	TMEM16H	86	16	0.317	FAILED
	GTPBP3	98	19	0.925	FAILED
Endogenous controls	β-actin SET 2	100	1	0.633	FAILED
	GAPDH SET 2	99	0	0.997	PASSED
	TPB SET 2	100	67	0.091	FAILED
	β-actin SET 3	100	0	0.979	PASSED
	GAPDH SET 3	99	0	0.994	PASSED
	TPB SET 3	98	20	0.707	FAILED
	β-actin SET 4	99	1	0.996	PASSED
	GAPDH SET 4	98	0	0.999	PASSED
	TPB SET 4	96	42	0.863	FAILED

Table 5.4**:** Quality control analysis performed for the Fluidigm expression data. Highlighted in red are the failed assays for a particular quality control criterion. 32 assays remained to be analysed after quality control. SD: Standard Deviation. (*):Assay passed because of an acceptable cell type specific pass rate.

5.3.3 Candidate gene expression analysis of the Fluidigm data and statistical methods used

The expression data were analysed as previously described using the comparative ΔΔCt method. As a calibrator for the analysis a randomly selected NOSE sample that exhibited 100% pass rate was used, NOSE11. Relative expression values of all cell lines to NOSE11 expression were generated and normalized against the average of GAPDH and β-actin for assays run in Fluidigm set 3 and 4 and only against GAPDH for assays run in Fluidigm set 2. Extreme outliers were removed when relative expression values were 10× more or less than the median of the relative values for each assay but it was checked that they would not exceed 5% of the samples**.**

Differences in the relative expression of each candidate gene between EOC cell lines and normal cell lines were assessed using the nonparametric two sided Wilcoxon Rank sum test using R software and P values were generated. Additionally, linear regression analysis was performed between NOSE, FTE and EOC cell lines. However these results could not be deemed as very informative since the number of FTE samples was too small (n=5) and will not be discussed. Statistically significant values were considered if $P<0.05$. However, in statistical analysis for many assays, genes can be found to be differentially expressed when they are not and these are deemed as false positives. Thus, it is vital to re-calculate the statistically significant p-value cut off when performing and statistical test on a big group of genes. The incidence of false positives is proportional to the number of tests performed. There are several multiple testing correction approaches but the more stringent, meaning allowing for the less false positive results to appear, is the Bonferroni correction. According to this, the statistically significant p-value of 0.05 is divided by the number of tests performed and the resulting value serves as the new significance level termed as p-value cut off according to which the results are deemed statistically significant or not. According to Bonferroni correction the statistically significant cut off value for the current analysis would be $P<0.05/32=0.0015$.

The differential expression of the endogenous controls that passed the QC analysis was compared between the different cell line sets as an additional

quality control criterion since their expression should not have been altered between EOC and normal cell lines. After correction for multiple testing none of the endogenous controls in any of the 3 Fluidigm chips run was found to be differentially expressed between EOC and normal cell lines further ensuring the reliability of the experiment.

Finally, it should be noted that the P values of differential expression were found to differ in some cases when normalised against β-actin or against GAPDH. The total number of assays that passed QC was 32 and out of those data were available for β-actin and GAPDH data for 25 assays. The concordance rate for the 25 assays when normalised against β-actin compared to GAPDH was 72% (18 out of 25). Assays where the P values were or were not statistically significant for both β-actin and GAPDH based on the P value calculated for Bonferroni correction were classed as being concordant. Out of the 7 assays that were not concordant the concordance difference found for 3 assays was because the comparison was based on the p value after Bonferroni correction but still showed the same trend of over or under expression. Because of these inconsistencies it was deemed important that the conclusions would be drawn from the p values obtained when the genes' expression normalised against the average expression of both the endogenous controls combined. Although the separate analyses are presented in the Table 1 of Appendix 3, P values presented in this chapter and the discussion will be based on the values obtained after the combined endogenous analysis with the exception of some genes that were run on Fluidigm chip 2 where only GAPDH passed the QC criteria.

The differential expression of selected candidate genes between EOC and normal (combined NOSE & FTE) cell lines in the Fluidigm assay was also compared with the differential expression of the genes using mRNA data for SOC and normal FT tissue from TCGA as previously described. There were 56 of the candidate genes with mRNA data available in TCGA. I treated the downloaded mRNA expression data as one experiment and have applied Bonferroni correction (cut-off $P< 8.93x10^{-4}$). Additionally, I investigated whether any of the candidate genes were found to have any somatic alterations. The query was performed in www.cbioportal.org which has the raw data of the

TCGA database. I have looked at the frequency of mutations, putative copy number (CN) alterations such as amplification or deletions. The sample size available for this analysis included 316 serous ovarian cancer samples. A summary describing the findings for analysed candidate genes is shown in Table 5.5.

5.3.4 Differential expression of candidate genes in locus 2q31 between normal and EOC cell lines- Extended study.

Twelve genes were investigated at locus 2q31 including *HOXD1* and *HOXD3* that were in closest proximity with SNP rs2072590 (Figure 5.11 A). Only 3 passed all the steps of QC and their differential expression was evaluated between EOC and normal cell lines (Table 5.5). Genes *HOXD1* and *HOXD3* did not pass the QC analysis. Thus the significant differential expression for *HOXD1* between NOSE and EOC that I previously identified could not be assessed in this experiment. According to the analysis performed using TCGA mRNA expression data these two genes did not seem to be differentially expressed in normal FT and SOC (Appendix 3, Table 1).

Of the three genes that passed all QC analysis, *KIAA1715* has been found to be under-expressed in EOC cell lines with a P value of 0.004. However, the result though did not remain significant after Bonferroni correction. Interestingly, analysis of TCGA data for this gene showed over-expression in SOC compared to normal FT with a P value of 1.89×10^{-4}, which remained significant after Bonferroni correction and was the most significant result amongst the genes tested in that locus according to the TCGA data. *MTX* was found to be over-expressed in EOC cell lines as well as in SOC compared to normal but neither result remained significant after Bonferroni correction (Table 5.5, Figure 5.11, B-E).

The selected genes were also evaluated for alterations such as amplifications, deletions, and somatic mutations. No more than 4.4% of the samples (n=316) was found to be altered in any of the selected genes with amplifications being the most frequent event as seen in Table 5.5.

			Fluidigm gene expression assay				TCGA mRNA epression analysis		TCGA analysis for somatic alterations	
Chromosome	Gene	Cytoband	Normalised against	Sample number (EOC/ Normal)	P value	Trend	P value	Trend	Number of samples with alterations * (Amp/ HomDel/ Mutations)	Frequency of alterations **
2	KIAA1715	2q31	β-actin & GAPDH	42/59	0.004	↓ EOC	1.89×10^{-4}	↑ SOC	11/0/0	3.5%
	HOXD8	2q31.1	β-actin & GAPDH	29/59	0.591	ND	0.245	ND	14/0/0	4.4%
	MTX2	2q31.1	β-actin & GAPDH	42/59	0.015	↑ EOC	0.004	↑ SOC	13/0/1	4.4%
3	SSR3	3q25.31	β-actin & GAPDH	42/58	0.003	↓ EOC	0.216	ND	22/0/0	7.0%
	TIPARP	3q25.31	β-actin & GAPDH	41/58	4.99×10^{-9}	↓ EOC	1.11×10^{-4}	↓ SOC	22/0/1	7.3%
	PA2G4P4	3q25.31	β-actin & GAPDH	41/61	2.14×10^{-5}	↑ EOC	No data		22/0/1	7.3%
8	MYC	8q24.21	β-actin & GAPDH	40/59	0.085	ND	0.059	ND	97/0/0	30.7%
	PVT1	8q24	β-actin & GAPDH	41/59	6.64×10^{-8}	↑ EOC	No data		102/0/0	32.3%
9	BNC2	9p22.32	GAPDH	30/58	1.73×10^{-4}	↓ EOC	1.63×10^{-4}	↓ SOC	8/2/1	3.5%
	CNTLN	9p22	GAPDH	42/58	0.030	↑ EOC	0.536	ND	4/3/0	2.2%
17	SP2	17q21.32	β-actin & GAPDH	42/58	5.68×10^{-10}	↑ EOC	0.049	↓ SOC	6/0/0	1.9%
	PNPO	17q21.32	β-actin & GAPDH	42/58	6.26×10^{-8}	↑ EOC	0.629	ND	6/0/0	1.9%
	CDK5RAP3	17q21.32	GAPDH	42/57	0.023	↑ EOC	0.143	ND	5/0/0	1.6%
	COPZ2	17q21.32	GAPDH	29/57	0.001	↓ EOC	0.107	ND	5/0/0	1.6%
	NFE2L1	17q21.3	β-actin & GAPDH	41/59	0.050	↓ EOC	1.4×10^{-5}	↓ SOC	5/0/0	1.6%
	CBX1	17q21.32	GAPDH	42/57	1.19×10^{-5}	↑ EOC	0.018	↑ SOC	6/0/0	1.9%
	SNX11	17q21.32	β-actin & GAPDH	42/58	0.004	↑ EOC	0.388	ND	5/0/0	1.6%
	HOXB3	17q21.3	β-actin & GAPDH	31/51	1.72×10^{-4}	↑ EOC	5.24×10^{-4}	↓ SOC	10/0/0	3.2%
	HOXB4	17q21-q22	β-actin & GAPDH	35/52	0.001	↑ EOC	0.026	↓ SOC	10/0/1	3.5%
	HOXB6	17q21.3	β-actin & GAPDH	26/50	4.23×10^{-7}	↑ EOC	5.25×10^{-5}	↓ SOC	8/0/0	2.5%
	HOXB7	17q21.3	β-actin & GAPDH	35/50	4.62×10^{-10}	↑ EOC	0.867	ND	8/0/0	2.5%
	HOXB9	17q21.3	β-actin & GAPDH	30/46	9.7×10^{-9}	↑ EOC	0.946	ND	8/0/0	2.5%
19	SIN3B	19p13.11	β-actin & GAPDH	42/57	0.126	ND	0.019	↑ SOC	28/0/2	9.5%
	HAUS8	19p13.11	β-actin & GAPDH	41/62	1.08×10^{-6}	↑ EOC	1.48×10^{-5}	↑ SOC	27/0/1	8.9%
	MYO9B	19p13.1	β-actin & GAPDH	42/59	0.010	↓ EOC	3.69×10^{-5}	↑ SOC	30/0/2	10.1%
	USE1	19p13.11	β-actin & GAPDH	43/60	1.69×10^{-5}	↑ EOC	0.224	ND	30/0/0	9.5%
	OCEL1	19p13.11	β-actin & GAPDH	38/59	0.012	↑ EOC	0.009	↓ SOC	30/0/0	9.5%
	NR2F6	19p13.1	β-actin & GAPDH	42/60	0.226	ND	0.094	ND	30/0/0	9.5%
	MERIT40	19p13.11	GAPDH	42/57	0.005	↑ EOC	2.09×10^{-5}	↑ SOC	29/0/1	9.5%
	ABHD8	19p13.11	GAPDH	42/58	0.028	↑ EOC	0.132	ND	27/0/0	8.5%
	DDA1	19p13.11	GAPDH	42/58	0.471	ND	2.04×10^{-4}	↑ SOC	27/0/0	8.5%
	MRPL34	19p13.1	β-actin & GAPDH	42/60	0.001	↑ EOC	0.005	↑ SOC	27/0/0	8.5%
	TMEM16H	19p13.11	β-actin & GAPDH	FAILED			7.42×10^{-5}	↑ SOC	26/0/3	9.2%
	GTPBP3	19p13.11	GAPDH	FAILED			1.02×10^{-5}	↑ SOC	25/0/0	7.9%

Table 5.5**:** Summary of the differential expression for candidate genes in the EOC associated loci. Additionally, tabulated is the differential expression evaluated for the same genes between SOC and normal FT tissue using TCGA expression data. Highlighted in yellow are statistically significant P values using the 0.05 cut-off. Highlighted in green are P values that remained statistically significant after Bonferroni correction, with a cut-off for the Fluidigm analysis of P<0.0015 and for the TCGA data analysis P<0.0008. The trend column indicates whether a gene was found to be over or under-expressed in EOC or SOC compared to normal cell lines or normal FT tissue respectively (ND: No Differential expression). Finally, identified somatic alterations for the candidate genes according to the TCGA are shown. Amp: amplification; HomDel: Homozygous Deletion; Mut: somatic mutation; (*) in a sample set of 316 SOC; (**) Alterations include Amp, HomDel and Mut.

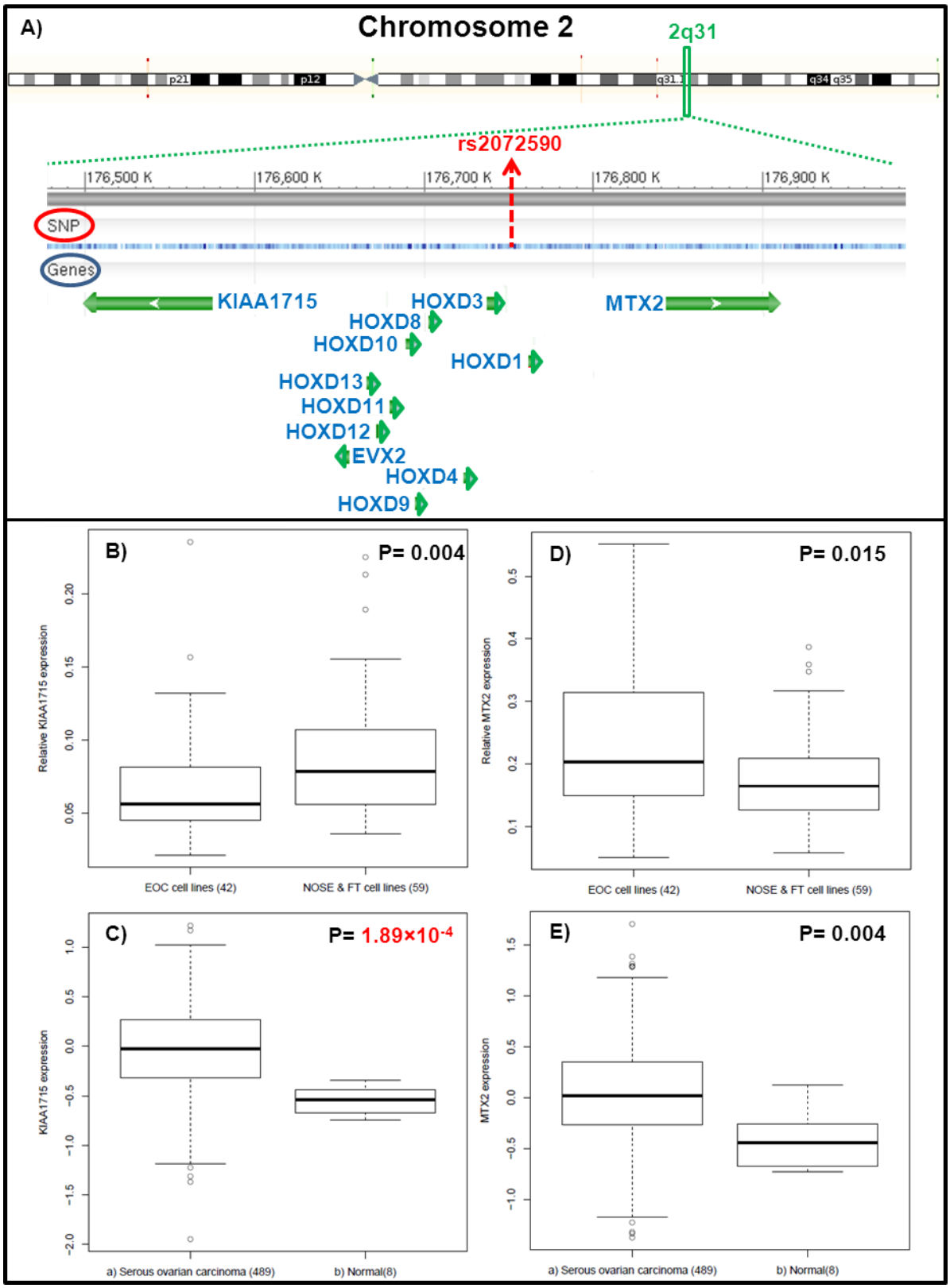


Figure 5.11**:** Mapping of SNP rs2072590 on chromosome 2 and expression analyses for candidate genes in locus 2q31. A) Location of candidate genes in locus 2q31. Differential expression between and EOC and normal cell lines for: B) *KIAA1715,* D) *MTX2.* Differential expression between SOC and normal FT for: C) *KIAA1715,* E) *MTX2.* P Values highlighted in red are the ones that remained significant after Bonferroni correction.

5.3.5 Differential expression of candidate genes in locus 3q25 between normal and EOC cell lines- Extended study

In locus 3q25, with *TIPARP* in the closest proximity to SNP rs2665390, five more genes were selected to be investigated (Figure 5.12 A). 3 of the 6 selected genes passed all the steps QC analysis and were further analysed for differential expression (Table 5.5).

TIPARP was found in the pilot study to be significantly under-expressed in EOC cell lines (Figure 5.5, $P< 10^{-8}$), suggesting a loss of function role for this gene in EOC development. This result was confirmed in this analysis in a higher number of normal and EOC samples ($P= 4.99\times10^{-9}$, Figure 5.12 B), consistent with the analysis preformed using TCGA data ($P= 1.11\times10^{-4}$, Figure 5.12 C). The pseudogene *PA2G4P4* was found to be over-expressed in EOC cell lines with a P value of 2.14×10^{-5} (Figure 5.12 F). This result could not be correlated to the expression in SOC and normal FT as there were no mRNA data available for this gene in TCGA. For gene *SSR3*, the expression was not found different in normal FT compared to serous carcinoma according to TCGA data. The slight under-expression of *SSR3* in EOC cell lines compared to normal cell lines with a P value of 0.003 did not remain significant after Bonferroni correction (Figure 5.12 D&E). *TIPARP* remains the most significant candidate gene within locus 3q25 according to these data.

A higher frequency of alterations was observed within the selected genes of chromosome 3 compared to chromosome 2. Somatic alterations in the genes for this locus ranged from 7 to 7.3% of the 316 SOC samples tested. All of the altered samples had amplifications for *SSR3* and only one of the 23 samples with alterations had a somatic mutations for *TIPARP* and *PA2G4P4* (Table 5.5) Interestingly, somatic alterations in *TIPARP* significantly decrease the protein levels of CDH2 according to reverse phase protein array (RPPA) data available from the TCGA ($P= 7.44E\times10^{-5}$) (www.cbioportal.org). CDH2 (Cadgerin-2) is a calcium dependent cell-cell adhesion glycoprotein and may be implicated in cancer metastasis when phosphorylated by the scr kinase pathway up-regulated by cancer cells.

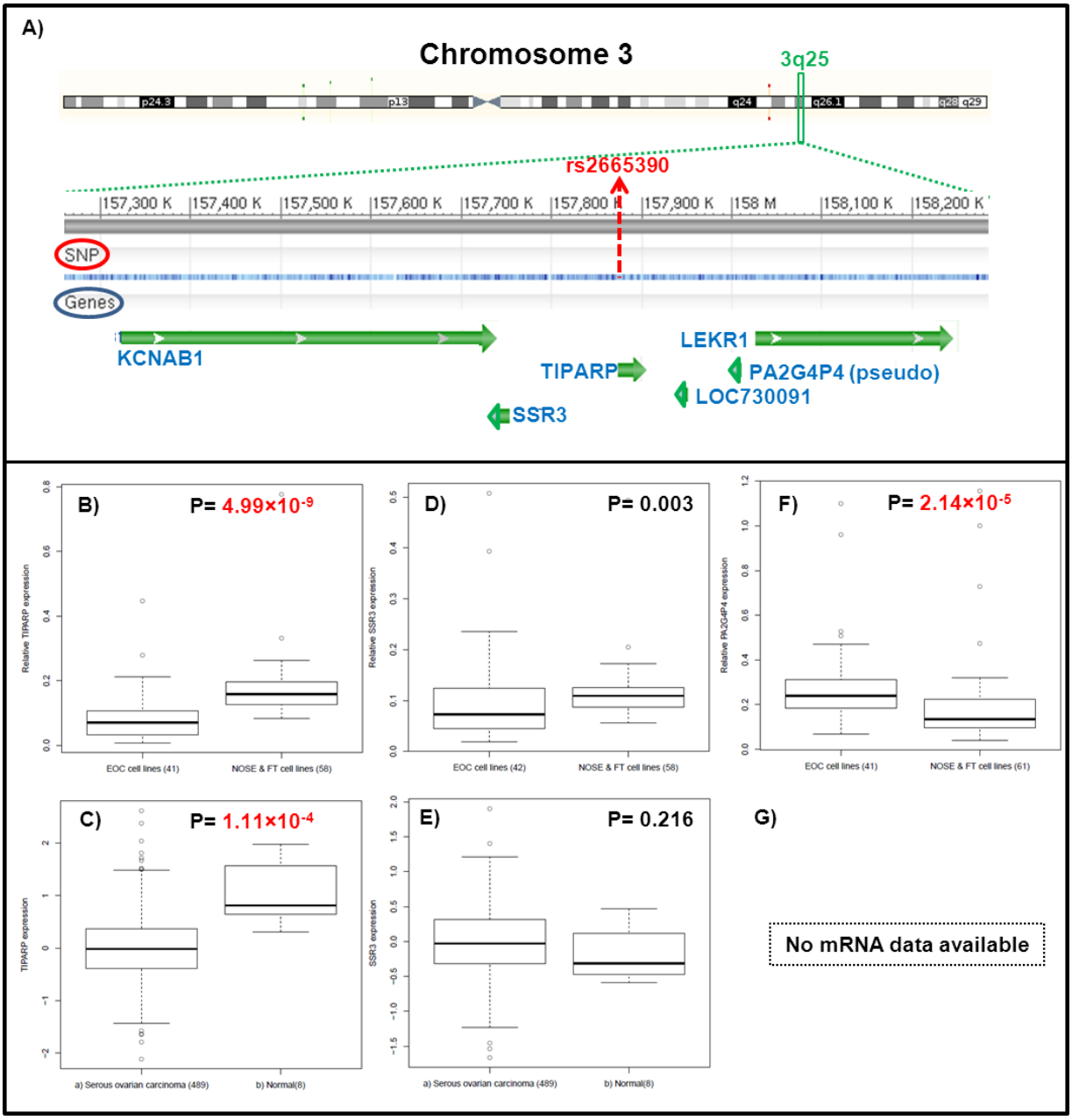


Figure 5.12**:** Mapping of SNP rs2665390 on chromosome 3 and expression analyses for candidate genes in locus 3q25. A) Location of candidate genes in locus 3q25. Differential expression between EOC and normal cell lines for: B) *TIPARP*, D) *SSR3* and F) *PA2G4P4*. Differential expression between SOC and normal FT for: C) *TIPARP*, E) *SSR3*. G) no mRNA expression data were available for *PA2G4P4* in TCGA. P Values highlighted in red are the ones that remained significant after Bonferroni correction.

5.3.6 Differential expression of candidate genes in locus 8q24 between normal and EOC cell lines- Extended study

Gene *PVT1* was selected to be investigated in addition to *MYC* at the 8q24 locus in proximity to SNP rs10088218 (Figure 5.13 A). This SNP is located in an apparent gene desert and these were the only genes in the 1Mb proximity of the SNP. Both of the selected genes passed all the steps of quality control analysis and were analysed for differential expression (Table 5.5)

The most interesting finding for this locus was *PVT1* which was found to be the extremely over-expressed in EOC cell lines suggesting a gain of function role for this gene in EOC development ($P= 6.64\times10^{-8}$, Figure 5.13 D). No mRNA data were available in TCGA for this gene.

MYC, an oncogene involved in the development of EOC and also found to be significantly over-expressed in EOC cell lines in the pilot study, was not confirmed to be over-expressed in the EOC cell lines in this extended study (P= 0.085, Figure 5.13 B). However, it is worth mentioning that the significance levels for *MYC* over-expression in the EOC cell lines found in the pilot study was low (ranging from P= 0.001 when normalised to β-actin to P= 0.02 when normalised against GAPDH). The expression of *MYC* was also not significantly different between normal FT and SOC (P= 0.059, Table 5.13, Figure 5.13 C).

A high frequency of amplifications was found for genes *MYC* and *PVT1* based on data from TCGA, reached a 32% frequency at the samples tested as seen in Table 5.5. All somatic alterations were due to gene amplifications. *PVT1* was found to have marginally more amplifications than *MYC*.

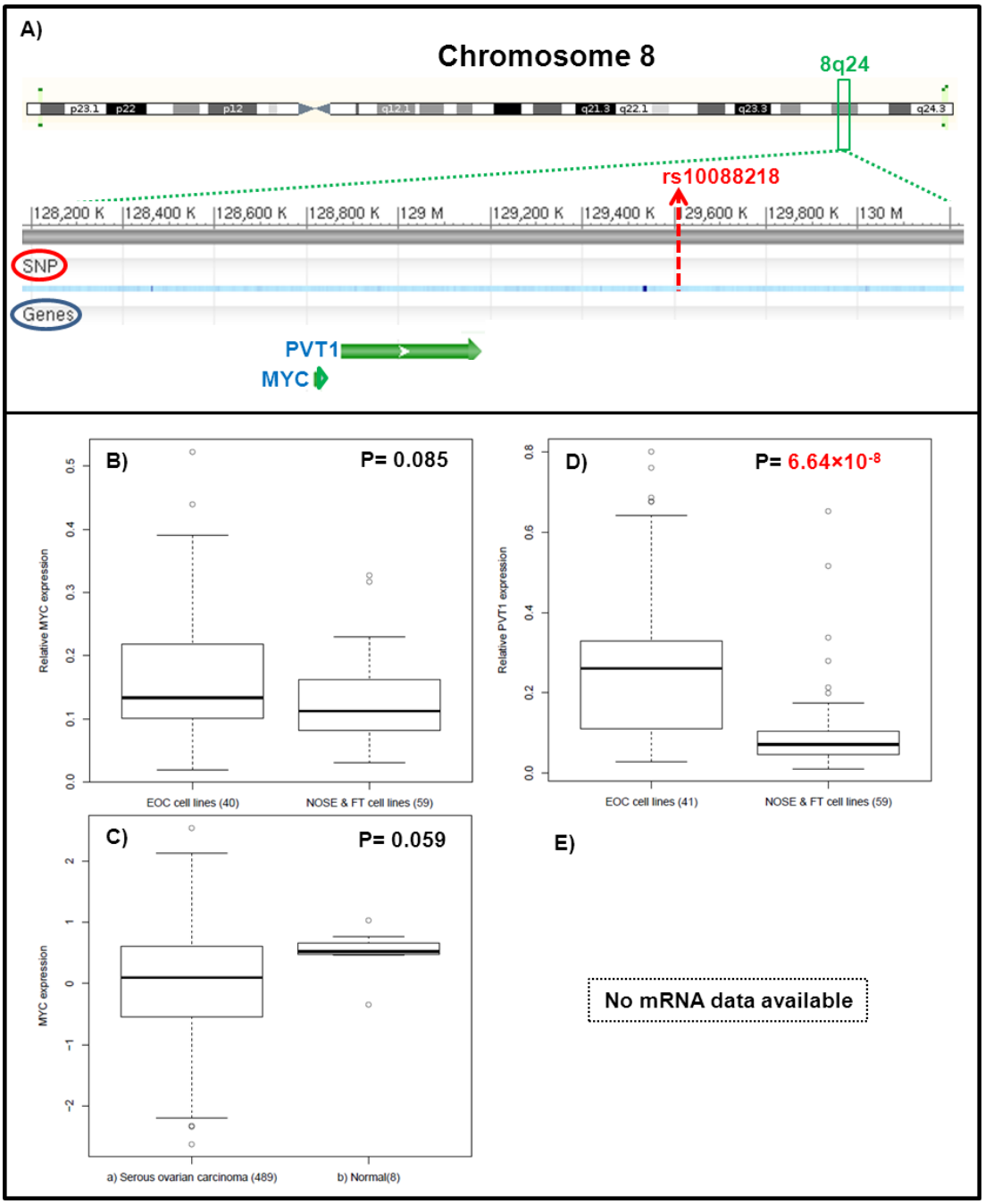


Figure 5.13**:** Mapping of SNP rs10088218 on chromosome 8 and expression analyses for candidate genes in locus 8q24. A) Location of candidate genes in locus 8q24. Differential expression between EOC and normal cell lines for: B) *MYC*, D) *PVT1*. Differential expression between SOC and normal FT for: C) *MYC,* E) *PVT1*. P Values highlighted in red are the ones that remained significant after Bonferroni correction.

5.3.7 Differential expression of candidate genes in locus 9p22 between normal and EOC cell lines- Extended study

In locus 9p22, with *BNC2* in the closest proximity to SNP rs3814113, one more gene, *CNTLN* was selected to be investigated (Figure 5.14 A). Both genes passed QC and their differential expression evaluated (Table 5.5). *BNC2* was previously found in the pilot Taqman gene expression assay to be extremely significantly under-expressed in EOC cell lines (Figure 5.7, $P<10^{-4}$), suggesting a loss of function role in EOC development. This result was confirmed in this analysis with a higher number of normal and EOC samples (Figure 5.14 B, $P= 1.73\times10^{-4}$), consistent with the analysis preformed using TCGA data (Figure 5.14 C, $P= 1.63\times10^{-4}$). The gene *CNTLN* was not significantly differentially expressed between EOC and NOSE cell lines after Bonferroni correction and similar results were found between SOC and normal FT (Figure 5.14, D&E). *BNC2* remains the most interesting candidate gene possibly implicated in EOC development within locus 9p22 based on the differential expression between EOC and normal cell lines.

A relatively small frequency of somatic alterations was observed within the selected genes of chromosome 9 with the highest frequency a 3.5% of tested samples with alterations within *BNC2*, attributed to amplifications, deletions and somatic mutations. Gene *CNTLN* was found to have equal amounts of amplifications and homozygous deletions in 2.2% of the samples tested (Table 5.5). I have previously mentioned that alterations in *TIPARP* affect CDH2 protein levels and it is interesting that alterations in *BNC2* affect the phosphorylation levels of the same protein according to reverse phase protein array (RPPA) data available from TCGA ($P= 7.44\times10^{-4}$) (www.cbioportal.org).

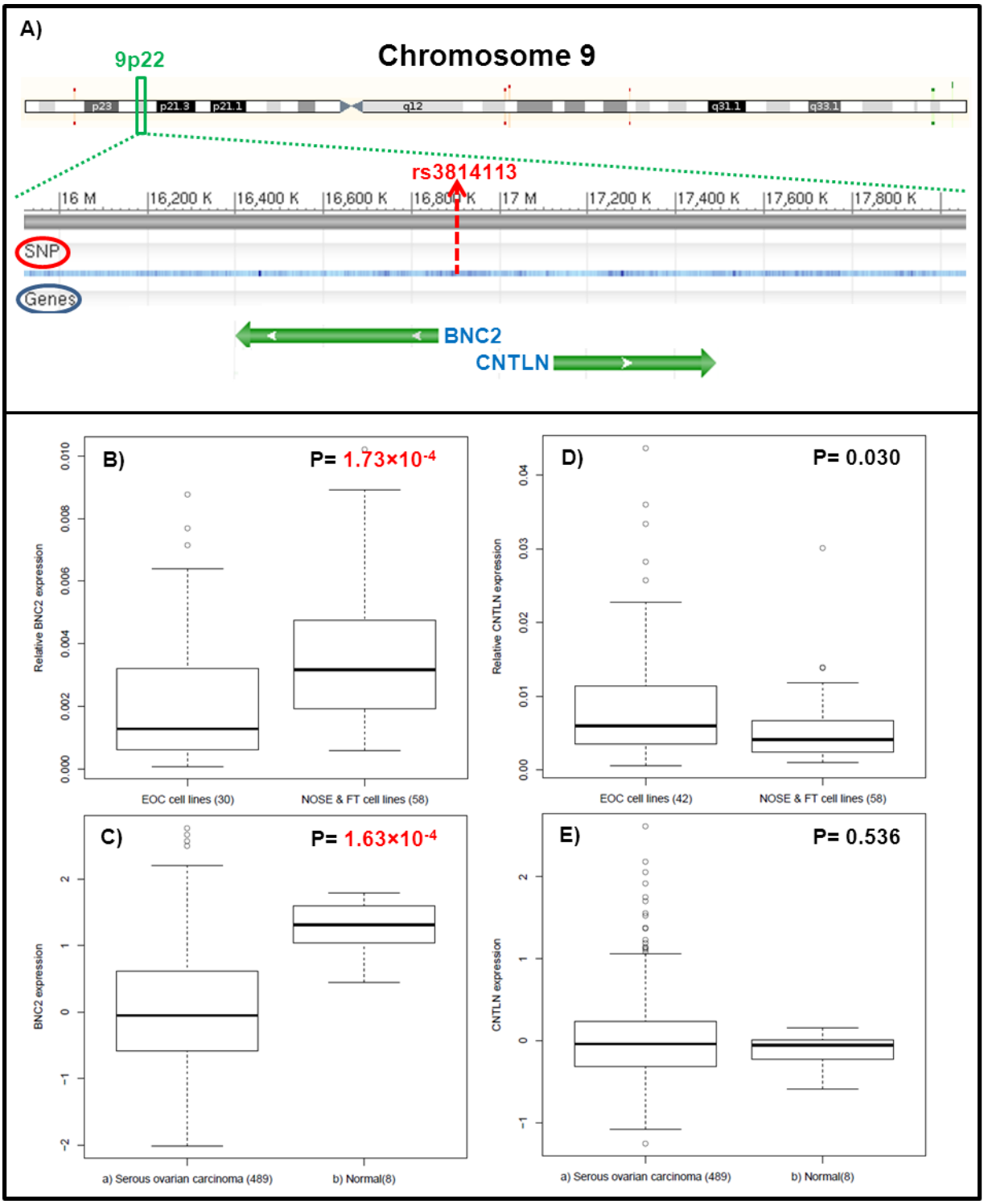


Figure 5.14**:** Mapping of SNP rs3814113 on chromosome 9 and expression analyses for candidate genes in locus 9p22. A) Location of candidate genes in locus 9p22. Differential expression between EOC and normal cell lines for: B) *BNC2*, D) *CNTLN*. Differential expression between SOC and normal FT for: C) *BNC2,* E) *CNTLN*. P Values highlighted in red are the ones that remained significant after Bonferroni correction.

5.3.8 Differential expression of candidate genes in locus 17q21 between normal and EOC cell lines- Extended study

Locus 17q21 is a gene rich region and a total of 23 genes spanning 1050Kb around the associated SNP were analysed (Figure 5.15 A). *SKAP1* was the gene in closest proximity with SNP rs9303542 since it is a *SKAP1* intronic SNP. However, *SKAP1* assay failed QC. Several other genes of this locus appeared to be differentially expressed between EOC & normal cell lines and between SOC and normal FT as shown in Table 5.5**.** Boxplots representing the significant findings for candidate genes from locus 17q21 in EOC versus normal cell lines and SOC versus normal FT are shown in Figure 5.15, B-U.

CBX1 may have a gain of function role in EOC as it was found significantly over-expressed in EOC compared to normal cell lines and remained significant after Bonferroni correction (P= 1.19×10^{-5}). Following the same trend, *CBX1* appeared over-expressed in SOC compared to normal FT (P= 0.018) but the result did not remain statistically significant after Bonferroni correction. *COPZ2* expression was significantly lower in EOC compared to normal cell lines after Bonferroni correction (P= 0.001), but the gene did not appear to be differentially expressed between SOC and normal FT. *NFE2L1* was not differentially expressed between EOC and normal cell lines but was significantly under-expressed in SOC compared to normal FT tissue (P= 1.4×10^{-5}).

HOXB3 expression was found to be significantly over-expressed in EOC compared to normal cell lines (P= 1.72×10^{-4}). Interestingly, *HOXB3* appeared to be significantly under-expressed in SOC compared to normal FT in contrast to the EOC versus normal cell lines model (P= 5.24×10^{-4}). A similar contradiction in expression trends between the two expression models was observed with the expression of *SKAP1* in the pilot study as well as *HOXB4* and *HOXB6* in this study that were found over-expressed in EOC cell lines but under expressed in SOC. *HOXB7, HOXB9, PNPO and SP2* were significantly over-expressed in EOC compared to normal cell lines (P= 4.64×10^{-10}, P= 9.7×10^{-9}, P= 6.26×10^{-8}, P= 5.68×10^{-10} respectively) but were not differentially expressed between SOC and normal FT.

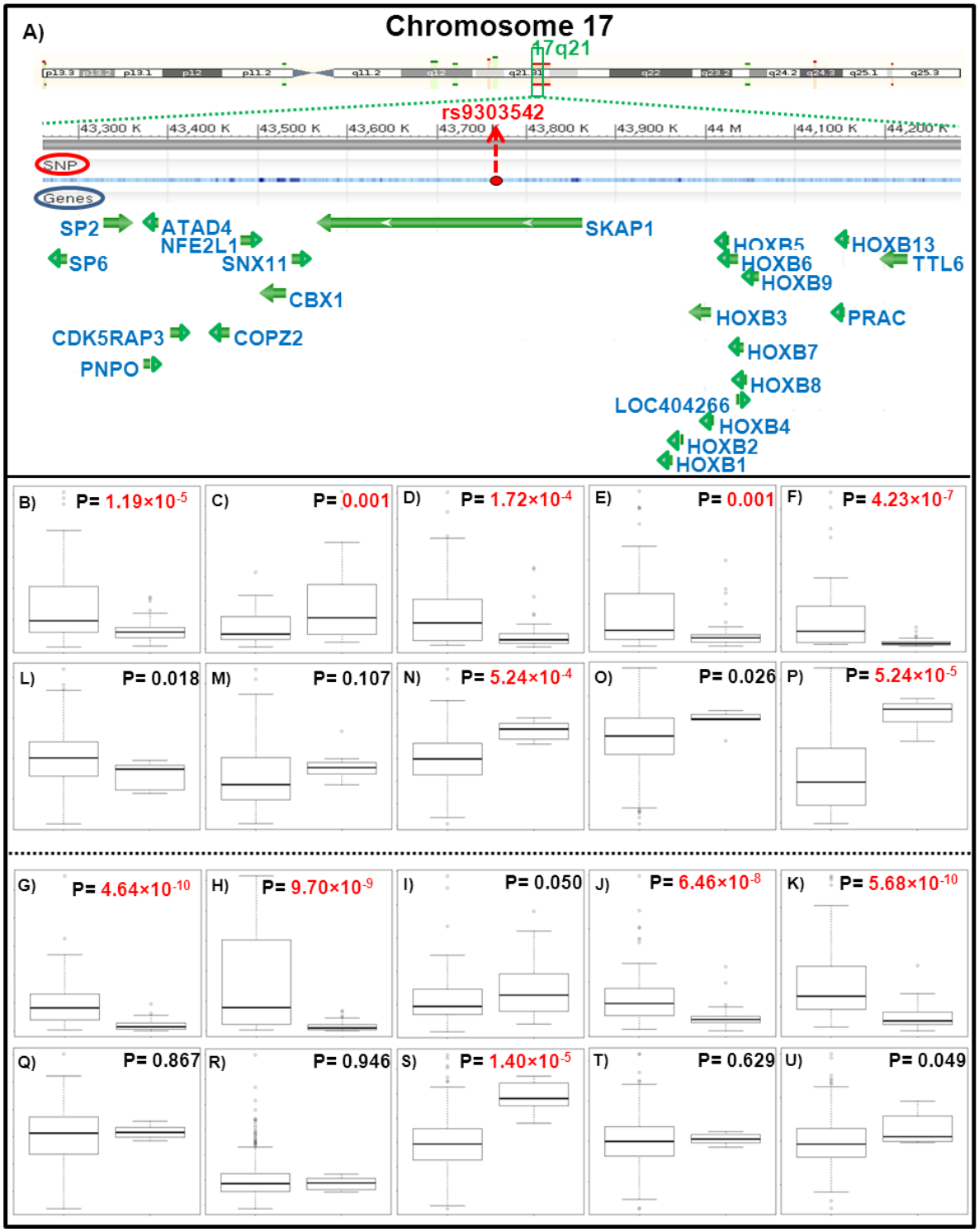


Figure 5.15**:** Mapping of SNP rs9303542 on chromosome 17 and expression analyses for candidate genes in locus 17q21. A) Location of candidate genes in locus 17q21. Differential expression between EOC (left hand side box) and normal cell lines (right hand side box) for: B) *CBX1*, C) *COPZ2*. D) *HOXB3*, E) *HOXB4*, F) *HOXB6*, G) *HOXB7*, H) *HOXB9*, I) *NFE2L1*, J) *PNPO*, K) *SP2*. Differential expression between SOC (left hand side box) and normal FT (right hand side box) for: L) *CBX1*, M) *COPZ2*. N) *HOXB3*, O) *HOXB4*, P) *HOXB6*, Q) *HOXB7*, R) *HOXB9*, S) *NFE2L1*, T) *PNPO*, U) *SP2*. P Values highlighted in red are the ones that remained significant after Bonferroni correction.

5.3.9 Differential expression of candidate genes in locus 19p13 between normal and EOC cell lines- Extended study

For chromosome locus 19p13, with *MERIT40* and *ANKRD41* in the closest proximity to SNPs rs8170 and rs2363956 respectively, 14 more genes were selected to be investigated (Figure 5.16 A). Five of the selected genes did not pass QC. The differential expression of the remaining nine genes was assessed and the results are shown in Table 5.5. Boxplots representing the significant differential expression found for candidate genes from locus 19p13 are shown in Figure 5.16, B-O.

MERIT40 was previously found in the pilot study to be significantly over-expressed in EOC cell lines (Figure 5.9, $P<10^{-5}$), suggesting a gain of function role in EOC development. This result was confirmed in this analysis with a higher number of normal and EOC samples ($P= 5\times10^{-3}$) consistent with the analysis preformed using TCGA data ($P= 2.09\times10^{-5}$). However, the EOC versus normal result did not remain significant after Bonferroni correction.

DDA1 and *MYO9B*(→after Bonferroni*)* expression were not found differentially expressed between EOC and normal cell lines but were significantly over-expressed in SOC compared to normal FT ($P= 2.09\times10^{-5}$ and $P= 3.69\times10^{-5}$ respectively). In contrast, the expression of *USE1* was significantly higher in EOC compared to normal cell lines ($P= 1.69\times10^{-5}$) but was not differentially expressed between SOC and normal FT tissue. *MRPL34* expression was significantly higher in EOC compared to normal cell lines (P= 0.001) consistent with a trend of *MRPL34* over-expression in SOC compared to normal FT (P= 0.005) but the latter not remain significant after Bonferroni correction. The most statistically significant finding for this locus was the *HAUS8* over-expression in EOC compared to normal cell lines ($P= 1.08\times10^{-6}$) consistent with the SOC and normal FT tissue expression model ($P= 1.48\times10^{-5}$). The assays for genes *TMEM16H* and *GTPBP3* did not pass QC but it is worth mentioning that they were both found significantly over-expressed in SOC compared to normal FT ($P= 7.42\times10^{-5}$ and $P= 1.02\times10^{-5}$ respectively) .

In conclusion, although at the 19p13 locus *HAUS8* is the gene most significantly associated with EOC development, *MERIT40* remains a very attractive candidate because of its functional link to BRCA1 involved in DNA

repair, and also because there is strong supporting evidence that the gene is frequently altered in ovarian tumour development. *MERIT40* was found to be amplified in ~9% of the SOC samples tested according to the TCGA (Table 5.5). TCGA data also suggest that that somatic alterations in *MERIT40* cause a significant decrease in the phosphorylation levels of the MAPK1/MAPK3 and FOXO3 proteins according to reverse phase protein array (RPPA) data available from TCGA ($P= 0.021$ and $P= 2.96\times10^{-4}$ respectively) as well as a decrease in MAPK9 protein levels ($P= 8.27\times10^{-4}$) which are proteins involved in the MAPK/ERK pathway which when deregulated can lead to uncontrolled cell growth and cancer development.

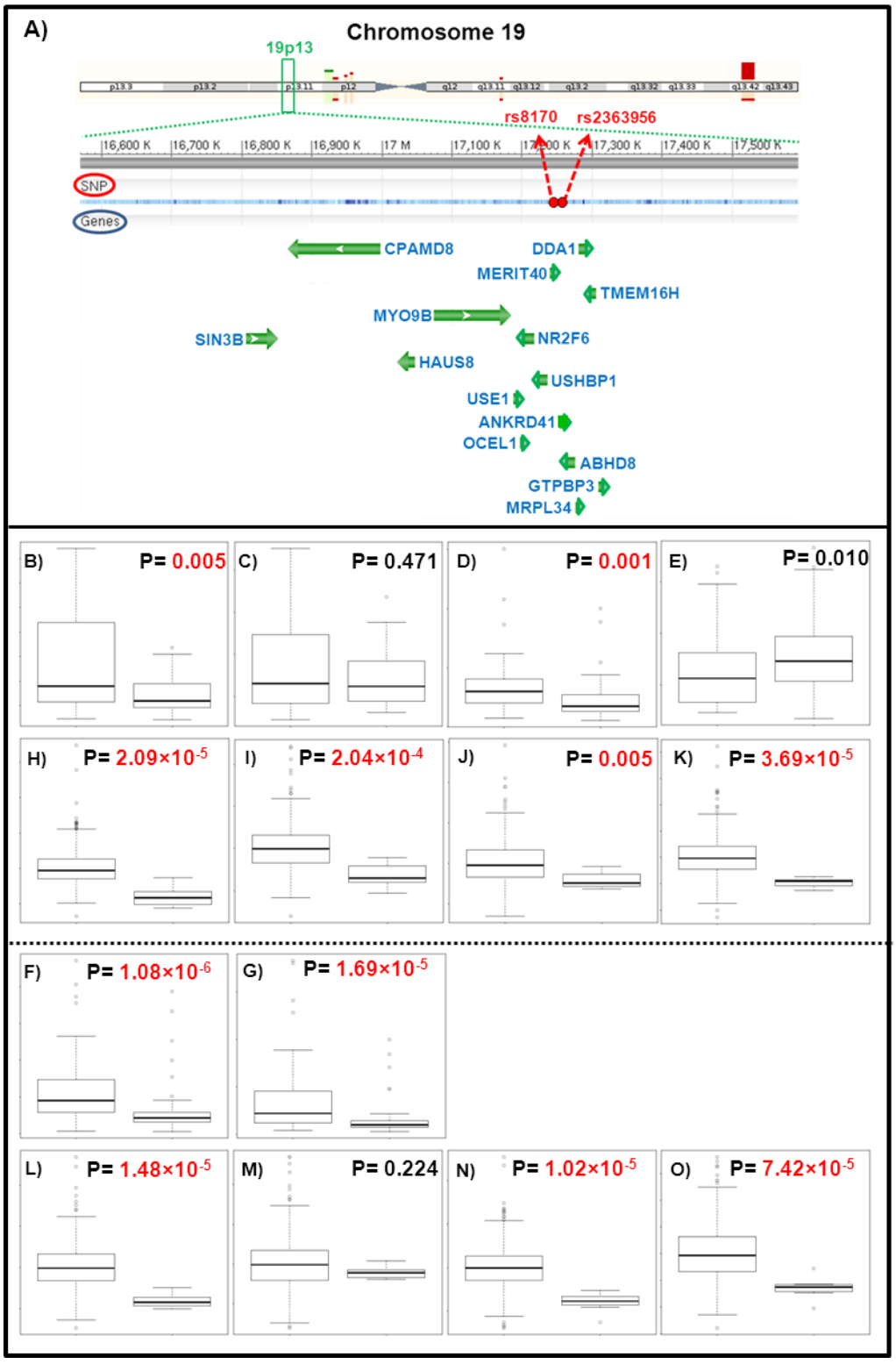


Figure 5.16**:** Mapping of SNPs rs8170 and rs2363956 on chromosome 19 and expression analyses for candidate genes selected to be investigated in locus 19p13. A) Location of candidate genes in locus 19q13. Differential expression between EOC (left hand side box) and normal cell lines (right hand side box) for: B) *MERIT40*, C) *DDA1*, D) *MRPL34*, E) *MYO9B,* F) *HAUS8,* G) *USE1*. Differential expression between SOC (left hand side box) and normal FT (right hand side box) for: H) *MERIT40*, I) *DDA1*, J) *MRPL34*, K) *MYO9B,* L) *HAUS8,* M) *USE1,* N) GTPBP3, O) TMEM16H. P Values highlighted in red are the ones that remained significant after Bonferroni correction.

5.4 Genotype specific expression of the candidate genes in the closest proximity to the EOC associated SNPs- Pilot Study

To test the functional link between the EOC susceptibility SNPs and gene expression I evaluated NOSE cell line expression for the candidate genes in close proximity with each SNP relatively to the genotype of the most significantly risk associated SNPs. DNA of 35 NOSE cell lines was genotyped for all the SNPs but due to the small sample size of this study I combined the expression data for rare homozygotes and heterozygotes and compared it to the expression of the common homozygotes. Gene expression analysis was performed as previously described using the comparative ΔΔCt method. Instead of using a single cell line as a calibrator sample, on this occasion the expression values of all cell lines were generated relative to the averaged Ct values of all common homozygotes per SNP. The normalization was performed against GAPDH and β-actin. Differences in the relative expression of each candidate gene between EOC cell lines and NOSE cell lines were assessed using the nonparametric two sided Wilcoxon Rank sum test using R software and P values were generated. Statistically significant values were considered if $P<0.05$. I found no evidence of genotype specific expression but the power of the study was limited because of the relatively small numbers. The results for genotype specific expression analysis are summarized on Table 5.6.

Gene	SNP	Relative expression of Candidate Genes for NOSE common homozygotes versus combined rare homozygotes & heterozygotes					
		Sample number CH/(RH+H)	Median of Relative expression		Median Absolute Deviation (MAD)		Wilcoxon Rank sum test (P value)
			CH	RH+H	CH	RH+H	
BNC2	rs3184113	20/27	0.93	1.00	0.27	0.21	0.975
SKAP1	rs9303542	18/17	0.92	0.94	0.35	0.39	0.858
TIPARP	rs2665390	31/4	0.94	1.37	0.27	0.14	0.049
CMYC	rs10088218	28/7	0.95	0.97	0.15	0.20	0.773
HOXD1	rs2072590	24/23	1.05	1.49	0.55	0.76	0.059
MERIT40	rs8170	23/12	1.00	1.01	0.19	0.23	0.797
	rs2363956	12/23	1.05	1.04	0.23	0.21	0.719

Table 5.6: Expression of the candidate genes relatively to NOSE cell lines' genotype of the most significant SNPs. CH: Common Homozygotes. RH+H: Rare Homozygotes & Heterozygotes combined.

5.5 Genotype specific expression of an extended list of candidate genes - Extended Study

Following the pilot study for genotype specific gene expression of the genes in closest proximity with the most significant SNPs of the GWAS no significant results were obtained. The extended analysis incorporated more normal (NOSE & FTE) cell lines increasing the power of this study and was performed for all the additional genes across the loci.

5.5.1 Quality control analysis

Quality control analysis was performed in order to guarantee the reliability of the gene expression data to be used for the genotype specific expression. All the assays that had a pass rate of <80%, based on the 64 normal samples on this occasion, were excluded from the analysis. 24 assays were excluded with a pass rate <80%. Five assays were excluded for either having a standard curve of R^2 <0.8 or because more than 15% of the samples demonstrated a standard deviation of >0.6. After all three steps of quality control analysis 31 assays remained to be analysed. A summary of the quality control analysis performed before the genotype specific gene expression analysis is shown on Table 5.7.

Of the 62 normal cell lines that remained after quality control analysis described before, 5 more NOSE and 1 more FTE sample were excluded from the analysis as there were no available genotyping data for these samples bringing the final number of samples to be analysed to 53 normal cell lines. Additionally, in order to assess the reliability of the genotyping data, the reported/expected minor allele frequency (MAF) from HAPMAP-CEU as found in the NCBI database was compared to the observed MAF from the genotyping of the NOSE and FTE samples (n=56). The formula that was used for this calculation was MAF= (2×r+h)/[2×(r+h+c), (where r= number of rare homozygotes, h= number of heterozygotes, c= number of common homozygotes). SNP rs2665390 located in chromosome 3 was not genotyped for the additional full set of NOSE cell lines but only for 33 in the pilot study. Therefore, SNP rs344008, which has an R^2=1 with rs2665390, that was

genotyped in the full set will also be used in the analysis. The expected and observed MAF was closely matched and is shown in Table 5.8.

Chromosome	Gene	Assay pass rate (%)	% of samples per assay with SD>0.6	R^2 Standard Curve	Assay failed or passed
2	KIAA1715	100	0	0.978	PASSED
	EVX2	0	0	FAILED	FAILED
	HOXD13	0	0	0.868	FAILED
	HOXD12	0	0	0.921	FAILED
	HOXD11	8	5	FAILED	FAILED
	HOXD10	41	14	FAILED	FAILED
	HOXD9	83	26	0.876	FAILED
	HOXD8	97	8	0.878	PASSED
	HOXD4	52	11	0.667	FAILED
	HOXD3	64	12	0.828	FAILED
	HOXD1	11	2	0.936	FAILED
	MTX2	100	0	0.805	PASSED
3	KCNAB1	70	8	FAILED	FAILED
	SSR3	100	0	0.993	PASSED
	TIPARP	100	0	0.936	PASSED
	LOC730091	52	15	0.052	FAILED
	PA2G4P4	100	0	0.956	PASSED
	LEKR1	39	9	0.016	FAILED
8	MYC	100	0	0.953	PASSED
	PVT1	97	17	0.834	FAILED
9	BNC2	100	2	0.980	PASSED
	CNTLN	100	0	0.944	PASSED
17	SP6	27	9	0.189	FAILED
	SP2	100	0	0.991	PASSED
	PNPO	100	0	0.955	PASSED
	ATAD4	0	0	0.991	FAILED
	CDK5RAP3	100	6	0.978	PASSED
	COPZ2	100	0	0.953	PASSED
	NFE2L1	100	0	0.953	PASSED
	CBX1	100	2	0.977	PASSED
	SNX11	100	0	0.963	PASSED
	SKAP1	11	2	0.559	FAILED
	HOXB1	0	0	FAILED	FAILED
	HOXB2	92	20	0.205	FAILED
	HOXB3	86	14	0.954	PASSED
	HOXB4	92	6	0.809	PASSED
	HOXB5	78	20	0.943	FAILED

	HOXB6	89	11	0.925	PASSED
	LOC404266	9	2	0.108	FAILED
	HOXB7	92	5	0.989	PASSED
	HOXB8	22	6	0.050	FAILED
	HOXB9	80	6	0.943	PASSED
	PRAC	0	0	FAILED	FAILED
	HOXB13	2	0	0.735	FAILED
	TTLL6	0	0	FAILED	FAILED
19	SIN3B	100	0	0.990	PASSED
	CPAMD8	0	0	0.653	FAILED
	HAUS8	100	2	0.982	PASSED
	MYO9B	100	0	0.983	PASSED
	USE1	100	2	0.906	PASSED
	OCEL1	100	6	0.890	PASSED
	NR2F6	100	2	0.944	PASSED
	USHBP1	9	9	FAILED	FAILED
	MERIT40	100	0	0.960	PASSED
	ANKRD41	75	23	0.717	FAILED
	ABHD8	100	0	0.922	PASSED
	DDA1	100	0	0.938	PASSED
	MRPL34	100	0	0.972	PASSED
	TMEM16H	89	20	0.317	FAILED
	GTPBP3	98	18	0.925	FAILED

Table 5.7**:** Quality control analysis performed for genotype specific expression for the Fluidigm assay. The quality control was performed based on the normal cell lines alone. 31 assays passed quality control.

Chromosome	SNP	Alleles (Major/Minor)	MAF HAPMAP CEU	MAF in normal cell lines
2	rs2072590	**G/T**	0.362	0.279
3	rs2665390	**T/C**	0.066	0.045
	rs344008 *	**C/T**	0.066	0.066
8	rs10088218	**G/A**	0.124	0.139
9	rs3814113	**T/C**	0.372	0.350
17	rs9303542	**A/G**	0.269	0.279
19	rs2363956	**T/G**	0.451	0.410
	rs8170	**C/T**	0.183	0.164

Table 5.8**:** HAPMAP-CEU MAF of EOC risk associated SNPs compared to the calculated NOSE and FTE cell line MAF. Highlighted in red are the minor alleles of the SNPs. MAF: Minor allele frequency. (*) R^2 =1 with rs2665390.

5.5.2 Genotype specific gene expression analysis and statistical tests performed- Extended study

The gene expression analysis was performed as previously described using the comparative ΔΔCt method. The expression values of all cell lines were generated relative to the averaged Ct values of all common homozygotes per SNP. The normalization was performed against GAPDH and/or β-actin depending on which Fluidigm chip the genes were run.

Differences in the relative expression of each candidate gene between EOC cell lines and NOSE cell lines were assessed using the nonparametric two sided Wilcoxon Rank sum test using R software and P values were generated. After Bonferroni correction for multiple testing (53 statistical tests throughout the 6 loci) the cut-off value for statistical significance was 0.0009. Following correction for multiple testing none of the results remained significant.

However, there were some genes that were significantly differentially expressed between common and the combined rare and heterozygote samples before Bonferroni correction as shown in Table 5.9. In chromosome 2, *HOXD1* was found to be over-expressed in rare homozygotes and heterozygotes compared to common homozygotes for SNP rs2072590 (P= 0.041). This finding could possibly suggest functional implication of the minor allele to the elevated expression of the gene (Figure 5.17 A). The most interesting finding of this analysis is the fact that the genotype of rs9303542 seems to be linked to the expression of three genes in chromosome 17. *CBX1* was over-expressed in the combined rare homozygotes and heterozygotes compared to the common homozygotes of this SNP (P value= 0.006, Figure 5.17 B). Both the expression of *SNX11* and *SP2* was higher in the common homozygotes suggesting a role of the minor allele of each SNP for down-regulation of these genes (Figure 5.17 C & D respectively) with a P value of 0.028 and 0.030 respectively. Increasing even further the number of normal samples might add more power to the study.

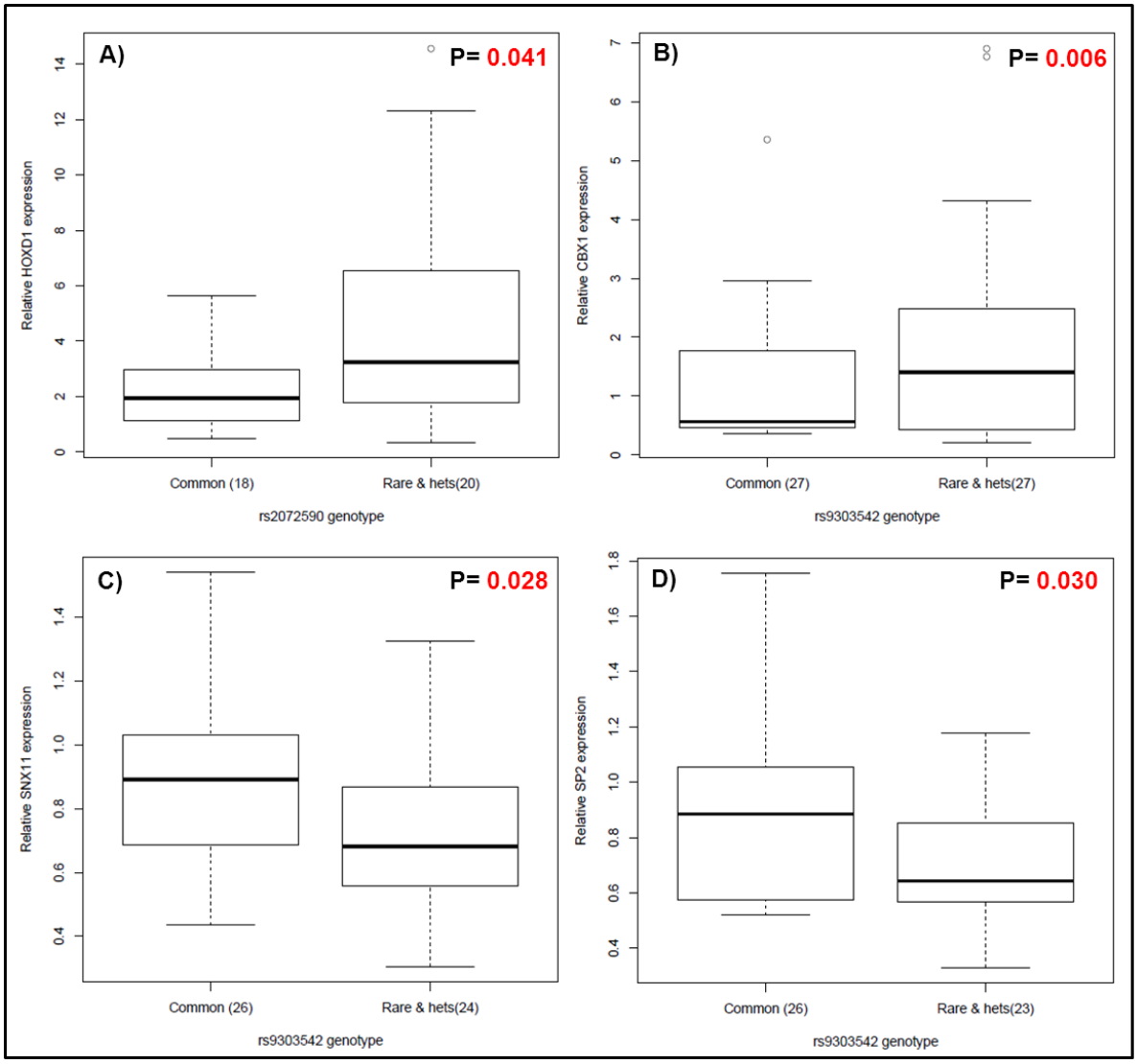


Figure 5.17**:** Boxplots of differential expression of candidate genes relatively to the genotype of EOC associated SNPs. A) *HOXD1* expression in normal cell lines relatively to rs2072590 genotype. B) *CBX1,* C) *SNX11* and D) *SP2* expression in normal cell lines relatively to rs9303542 genotype. Common: common homozygotes; Rare: rare homozygotes; Hets: heterozygotes.

Chromosome	SNP	Genotype (Major/Minor)	Gene	Relative expression of common homozygotes compared to combined rare homozygotes and heterozygotes					
				Sample Number CH/(RH+H)	Median of Relative expression		Median Absolute Deviation (MAD)		Wilcoxon Rank sum test (P value)
					CH	RH+H	CH	RH+H	
2	rs2072590	G/T	KIAA1715	22/28	0.98	1.18	0.35	0.32	0.516
			HOXD8	21/30	1.31	1.43	0.43	0.67	0.629
			HOXD1 (*)	18/20	1.95	3.23	0.91	1.98	**0.041**
			MTX2	23/28	0.76	0.77	0.21	0.22	0.659
3	rs2665390	T/C	SSR3	23/3	0.66	0.63	0.17	0.10	0.940
			TIPARP	23/3	0.64	0.79	0.20	0.13	0.395
			PA2G4P4	23/3	0.64	0.80	0.15	0.11	0.352
	rs344008	C/T	SSR3	42/7	0.72	0.73	0.15	0.04	0.685
			TIPARP	42/7	0.71	0.90	0.17	0.10	0.179
			PA2G4P4	43/7	0.72	0.87	0.25	0.20	0.295
8	rs10088218	G/A	MYC	39/12	0.77	0.76	0.28	0.21	0.887
9	rs3814113	T/C	BNC2	22/27	0.87	0.73	0.34	0.29	0.353
			CNTLN	21/27	0.77	0.70	0.37	0.27	0.313
17	rs9303542	A/G	SP2	26/23	0.89	0.64	0.22	0.13	**0.030**
			PNPO	26/24	0.84	0.71	0.17	0.21	0.079
			CDK5RAP3	25/24	0.82	0.66	0.46	0.32	0.481
			COPZ2	26/24	0.86	0.41	0.50	0.18	0.360
			NFE2L1	26/25	0.81	0.86	0.24	0.36	0.772
			CBX1	25/24	0.70	0.43	0.19	0.13	**0.006**
			SNX11	26/24	0.89	0.68	0.16	0.15	**0.028**
			SKAP1 (*)	22/17	0.88	0.87	0.32	0.42	0.510
			HOXB3	19/23	0.86	0.94	0.53	0.54	0.653
			HOXB4	20/22	1.20	1.73	0.56	0.60	0.071
			HOXB6	20/24	1.12	1.47	0.44	0.68	0.420
			HOXB7	21/22	1.38	1.29	1.04	0.62	0.782
			HOXB9	17/23	0.71	0.87	0.49	0.52	0.705
19	rs2363956	T/G	SIN3B	18/31	1.14	1.12	0.27	0.28	0.382
			HAUS8	18/36	1.27	1.51	0.65	0.37	0.105
			MYO9B	18/33	0.99	1.12	0.23	0.27	0.395
			USE1	18/32	1.02	1.10	0.27	0.37	0.405
			OCEL1	18/32	0.96	0.96	0.38	0.44	0.772
			NR2F6	18/34	1.04	0.98	0.36	0.37	0.977
			MERIT40	18/32	0.87	0.91	0.42	0.26	0.881
			ABHD8	18/32	1.02	0.93	0.67	0.41	0.865

			DDA1	18/32	0.87	1.07	0.32	0.40	0.555
			MRPL34	18/34	1.01	1.20	0.27	0.47	0.132
	rs8170	C/**T**	SIN3B	37/16	0.78	0.75	0.19	0.24	0.752
			HAUS8	35/17	0.96	1.17	0.50	0.63	0.802
			MYO9B	37/16	0.88	0.72	0.37	0.44	0.425
			USE1	35/14	0.76	0.68	0.20	0.24	0.577
			OCEL1	38/16	0.89	0.61	0.38	0.26	0.076
			NR2F6	38/17	1.04	0.61	0.70	0.41	0.290
			MERIT40	36/15	0.70	0.71	0.43	0.39	0.927
			ABHD8	34/15	0.71	0.62	0.44	0.35	0.974
			DDA1	36/15	0.68	0.67	0.40	0.37	0.862
			MRPL34	37/17	0.87	0.61	0.59	0.50	0.605

Table 5.9: Summary of genotype specific gene expression analysis for the extended list of candidate genes in the six risk loci. Tabulated are relative expression median values for the candidate genes compared between common homozygotes and combined rare homozygotes and heterozygotes. Highlighted in blue are the candidate genes initially selected with the closest proximity to the most significant SNPs in each locus. Highlighted in yellow are P values for assays that were statistically significant before application of multiple testing correction cut-off. (*) Genotype specific expression from those genes was assayed using the expression data from the pilot study as they failed QC in the extended study. CH: Common Homozygotes. RH+H: Rare Homozygotes + Heterozygotes combined.

5.6 Investigating genotype specific methylation for CpGs in the candidate genes- Pilot study

I decided to investigate whether the EOC risk SNPs have a functional role in conferring disease susceptibility by regulating the methylation status of the candidate genes in healthy individuals. Within our unit a methylation analysis has been performed for approximately 27,000 CpGs mapping to promoters of 14,000 genes in germline DNA from 148 healthy controls (Teschendorff et al, 2009) that were also included as part of Phase II of our GWAS and had been genotyped for the SNPs of interest. A limitation of this study is that methylation is known to be tissue specific and it would be desirable to have methylation data from the cells of origin for EOC.

I performed an analysis comparing the methylation status in CpG islands of the selected candidate genes and the different genotype groups of the significant SNPs. There were no methylation data available for CpGs in *MERIT40* and *SKAP1* gene promoters. Initially, linear regression analysis was performed in order to compare methylation status between the 3 genotype groups; rare and common homozygotes and heterozygotes. However, because the sample size was small the rare homozygotes were further combined with the heterozygotes and compared to the common homozygotes. Wilcoxon rank sum test was then performed comparing the methylation status of the common and the combined heterozygotes & rare homozygotes. I found marginally significant associations between hypomethylation of CpG island cg19972619 of *MYC* and the common homozygote genotypes of SNP rs10088218 (P= 0.047). I also found association of hypomethylation of cg19001226 near the *HOXD1* gene with the common homozygote genotype of rs2072590 (P= 0.035). The results did not remain significant after Bonferroni correction (cut-off P value for statistical significance: 0.006 based on 8 statistical tests performed). Table 5.10 summarises the findings of this analysis for all candidate SNPs. The box plots for significant methylation-genotype associations are shown in Figure 5.18

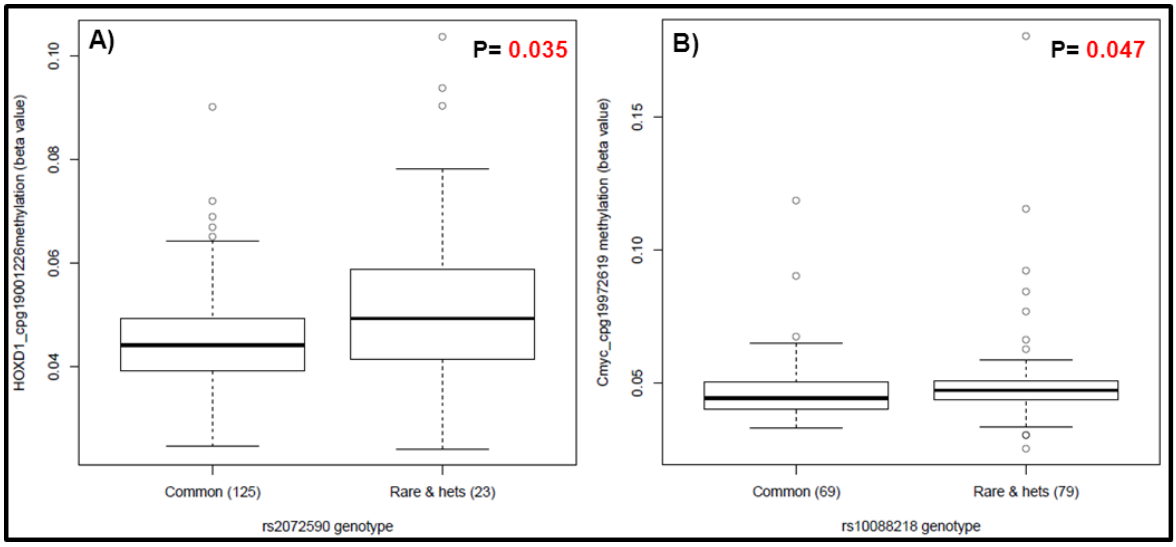


Figure 5.18**:** Genotype specific methylation analysis for CpG sites of *HOXD1* and *MYC* genes. (A) Methylation data for cg19001226 of *HOXD1* indicate that individuals carrying the common GG genotype of rs2072590 were hypomethylated compared to the combined rare homozygotes and heterozygotes. (B) Methylation data for cg19972619of *MYC* indicate that individuals carrying the common homozygous genotype of rs10088218 were hypomethylated.

Chromosome	SNP	Genotype (Major/Minor)	Gene	CpG site	Sample number CH/(RH+H)	median β-value comparison between common homozygotes and combined rare homozygotes & heterozygotes				
						Median of β-values		Median Absolute Deviation (MAD)		Wilcoxon Rank sum test(P value)
						CH	RH+H	CH	RH+H	
2	rs2072590	G/T	HOXD1	cg19001226	125/23	0.04	0.049	0.01	0.008	**0.035**
			HOXD3	cg00005847		0.16	0.160	0.02	0.016	0.992
				cg18702197		0.06	0.059	0.01	0.009	0.075
3	rs2665390	T/C	TIPARP	cg22114222	125/23	0.02	0.020	0.00	0.002	0.136
				cg24262469		0.21	0.204	0.03	0.035	0.866
8	rs10088218	G/A	MYC	cg19972619	69/79	0.04	0.047	0.01	0.004	**0.047**
				cg27207274		0.07	0.071	0.01	0.006	0.604
9	rs3184113	T/C	BNC2	cg24341129	61/87	0.28	0.287	0.04	0.031	0.852

Table 5.10**:** Summary of genotype specific methylation analysis of candidate genes in loci 2q31, 3q25, 8q24 and 9p22 for the EOC risk associated SNPs (Pilot study). Tabulated are the CpGs' median methylation (represented by β-values) of common homozygotes compared to the median methylation of combined rare homozygotes and heterozygotes. Highlighted in yellow are P values that are significant before Bonferroni correction. CH: Common Homozygotes. RH+H: Rare Homozygotes + Heterozygotes.

Many studies have shown that methylation status correlates with the age of the patients. Therefore, to ensure that the different age groups did not show differential methylation, which might have been responsible for the SNP associations with methylation I found, I performed linear regression analysis for methylation status across 3 age groups 50-59, 60-69 and >70 (Table 5.11). The methylation of the analysed CpGs was not affected by age. Thus, none of the positive associations I found were biased by age.

Gene	CpG island	Linear regression P value_ Age affecting methylation
HOXD1	cg19001226	0.062
HOXD3	cg00005847	0.567
	cg18702197	0.369
TiPARP	cg22114222	0.950
	cg24262469	0.209
Cmyc	cg19972619	0.180
	cg27207274	0.935
BNC2	cg24341129	0.236

Table 5.11: Linear regression analysis of candidate gene methylation based on 3 age groups. Group 1: 50-59 yrs, Group 2: 60-69 yrs, Group 3: >70 yrs. No significant association of age with methylation was found for the CpGs associated with the candidate genes.

5.7 Genotype specific methylation for CpGs in the extended group of candidate genes- Extended study

I previously performed an analysis comparing methylation status of CpG islands in the six genes in closest proximity to the significant SNPs across the different genotype groups for the most significant SNPs identified by the GWAS. This analysis was performed using methylation and genotyping data from 148 healthy controls. In that initial study I found significant associations between genotypes and methylation status of *MYC* CpG island cg19972619 for SNP rs10088218. Also methylation in cg19001226 of *HOXD1* gene was found to be significantly associated with rs2072590. However none of the results remained significant after Bonferroni correction. Increasing the sample size of this study could increase the power to detect stronger positive associations.

Therefore, I have expanded this analysis by increasing the sample size to 256 healthy controls with an additional 108 healthy individuals available with methylation data (Teschendorff et al, 2010). Additionally, I evaluated genotype specific methylation for CpG islands in 47 of the candidate genes across the 6 loci for which methylation data were available. In the initial analysis, because the sample size was smaller (n=148), the number of rare homozygotes was combined with the heterozygotes that contain the rare allele and Wilcoxon rank sum test analysis was performed between the common compared with the heterozygotes and rare combined. For the extended analysis, in order to compare with the pilot study, I also performed Wilcoxon rank sum test to investigate the methylation status of the combined rare homozygotes and heterozygotes compared to the common homozygotes' methylation status. However, because the sample size was larger (methylation data from peripheral blood DNA from 256 unaffected individuals compared to 148 previously used in the pilot study) I also performed linear regression analysis in order to compare methylation status between the 3 genotype groups, rare and common homozygotes and heterozygotes (Appendix 3, Tables 2-4).

In the present analysis the statistical significant cut-off P-value was 0.0007 after Bonferroni correction for multiple testing (statistical tests performed=71). A summary of the genotype specific methylation analysis, using the Wilcoxon rank sum test analysis, performed for the extended list of genes is

presented in Table 5.12 for the risk loci of chromosomes 2, 3, 8, 9 and in Table 5.13 for the risk loci of chromosomes 17 and 19. The associations found in the pilot study did not remain after analysis in the bigger sample size. Significant associations were found between methylation status of CpG island cg01405107 of *HOXB5* and genotype of SNP rs9303542 and locus 17q21 (P= 2.11×10^{-5}, Figure 5.18 A). In more detail, individuals carrying the common AA genotype of SNPs rs9303542 were hypermethylated at the CpG island cg01405107 of *HOXB5* compared to the heterozygous and rare homozygous individuals. The association was also significant when linear regression was performed comparing the methylation status across the 3 different genotypes of this SNP (P= 5.56×10^{-6}, Appendix 3, Table 3). It is interesting that methylation of cg01405107 reduced relatively to the presence of the minor allele with a distinct trend of the highest methylation in common homozygotes containing no copies of the minor allele, less methylation for the heterozygotes that contain one copy of the minor allele and the lowest level of methylation for the rare homozygotes containing two copies of the minor allele (Figure 5.19 B).

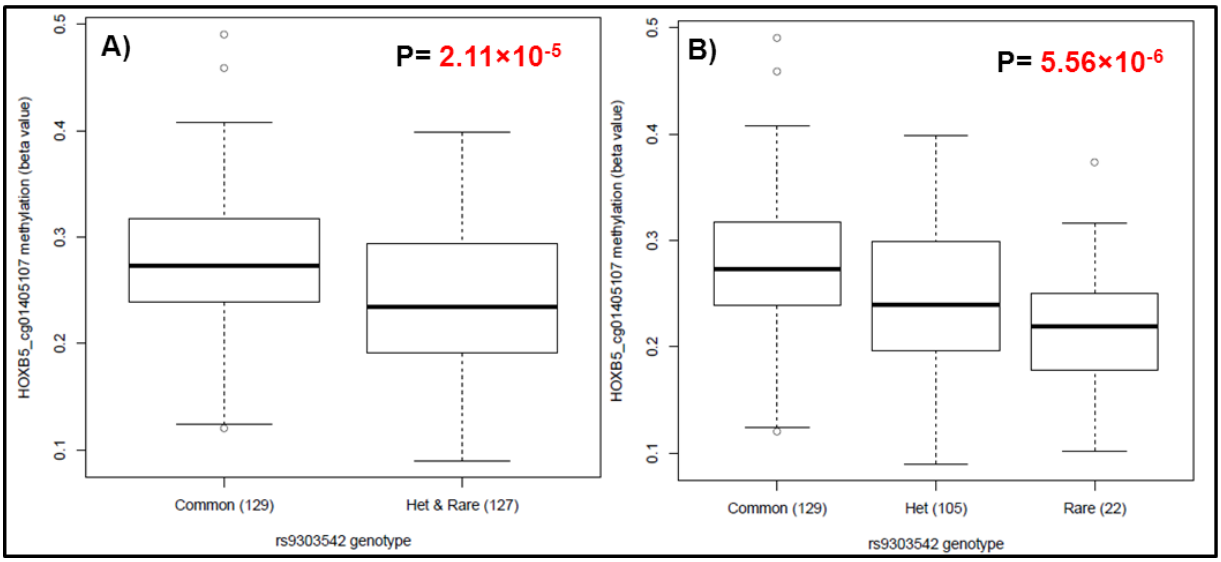


Figure 5.19**:** Genotype specific methylation analysis for cg01405107 associated with *HOXB5* gene relative to the genotype of SNP rs9303542. A) Differential methylation of cg01405107 of *HOXB5* between common homozygotes and the combined rare homozygotes & heterozygotes of rs9303542. Wilcoxon rank sum test performed for the P value shown B) Differential methylation of cg01405107 of HOXB5 between common homozygotes and rare homozygotes and heterozygotes. Linear regression analysis was performed for the P value shown.

I also found other weak associations that did not remain significant after Bonferroni correction. In chromosome 17q21 associations were found between methylation status of cg16848873 in *HOXB6*, cg06760035 in *HOXB4* and rs9303542 genotype (P= 0.005 and P= 0.031 respectively). Another association was observed between methylation status of cg10957151 in *HOXD9* and genotype of SNP rs2072590 in 2q31 (P= 0.001). Finally, in locus 19p13, I found associations between methylation status of cg21435336 in *ANKRD41*, cg22734480 in *ABHD8* and rs2363956 genotype (P= 0.004 and P= 0.038 respectively). All the associations were also significant in linear regression analysis performed comparing the methylation between the three genotype groups (Appendix 3, Tables 2, 3, ,4). I found no association of methylation status of any CpG islands with SNP genotypes for any genes in the risk loci of chromosomes 3, 8, and 9.

Finally, I investigated whether the age of the individuals would be associated with methylation status of the studied CpGs and thus driving the genotype specific methylation results. Similarly to the pilot study, I performed linear regression analysis for methylation status across 3 age groups 50-59, 60-69 and >70. A summary of the linear regression analysis and P values generated to investigate if age affects methylation of the investigated CpGs is presented in Table 5.14. The methylation of the analysed CpGs was not significantly affected by age when the Bonferroni cut-off P value was used to evaluate the results. However, some weak associations were found between age and methylation status of CpGs associated with genes *HOXD1, HOXD9, HOXD8, PVT1, SKAP1, HOXB2* and *HOXB4* before Bonferroni correction. These results indicate that the most significant association found between *HOXB5* CpG methylation and rs9303542 was not biased by age.

Chromosome	SNP	Genotype (Major/Minor)	Gene	CpG site	Median β-value comparison between common homozygotes and combined rare homozygotes & heterozygotes				
					Sample number CH/(RH+H)	Median of β-values		Median Absolute Deviation (MAD)		Wilcoxon Rank sum test (P value)
						CH	RH+H	CH	RH+H	
2	rs2072590	G/T	HOXD9	cg10957151	128/128	0.174	0.157	0.020	0.022	0.001
				cg14991487		0.053	0.050	0.007	0.009	0.393
			HOXD8	cg15520279		0.036	0.033	0.009	0.008	0.350
				cg21815667		0.074	0.075	0.009	0.007	0.863
			HOXD3	cg00005847		0.150	0.151	0.017	0.015	0.509
				cg18702197		0.053	0.054	0.007	0.008	0.663
			HOXD1	cg19001226		0.041	0.042	0.006	0.006	0.730
			MTX2	cg19280014		0.083	0.084	0.008	0.008	0.353
3	rs2665390	T/C	KCNAB1	cg15423862	210/46	0.855	0.856	0.010	0.012	0.328
				cg23873703		0.695	0.703	0.058	0.062	0.696
			SSR3	cg24517609		0.073	0.071	0.006	0.007	0.936
			TIPARP	cg22114222		0.022	0.021	0.002	0.002	0.240
				cg24262469		0.196	0.208	0.022	0.023	0.199
8	rs10088218	G/A	MYC	cg19972619	198/58	0.044	0.044	0.004	0.004	0.769
				cg27207274		0.071	0.072	0.006	0.006	0.178
			PVT1	cg13052755		0.886	0.886	0.014	0.011	0.428
				cg13784855		0.038	0.037	0.003	0.003	0.785
9	rs3814113	T/C	BNC2	cg24341129	108/148	0.276	0.282	0.035	0.032	0.366

Table 5.12**:** Summary of genotype specific methylation of CpG islands associated with the extended list of candidate genes within loci in chromosomes 2, 3, 8 and 9. Tabulated are the CpGs' median methylation (represented by β-values) of common homozygotes compared to the median methylation of combined rare homozygotes and heterozygotes. P values were generated using the Wilcoxon rank sum test. Highlighted in yellow are P values that are significant before Bonferroni correction. CH: Common Homozygotes. RH+H: Rare Homozygotes + Heterozygotes combined.

Chromosome	SNP	Genotype (Major/Minor)	Gene	CpG site	Median β-value comparison between common homozygotes and combined rare homozygotes & heterozygotes					
					Sample number CH/(RH+H)	Median of β-values		Median Absolute Deviation (MAD)		Wilcoxon Rank sum test (P value)
						CH	RH+H	CH	RH+H	
17	rs9303542	A/G	SP2	cg02691091	129/127	0.906	0.904	0.007	0.007	0.352
				cg14360917		0.530	0.517	0.040	0.027	0.057
			PNPO	cg00177698		0.037	0.036	0.004	0.004	0.085
				cg10154655		0.041	0.042	0.004	0.004	0.283
			NFE2L1	cg03777405		0.062	0.064	0.008	0.007	0.313
				cg06740897		0.083	0.086	0.012	0.015	0.745
			CBX1	cg11194725		0.053	0.052	0.005	0.004	0.229
				cg17778721		0.040	0.040	0.006	0.006	0.620
			SNX11	cg08413153		0.111	0.116	0.012	0.013	0.207
				cg12563480		0.035	0.036	0.006	0.006	0.667
			SKAP1	cg05106502		0.368	0.378	0.039	0.034	0.439
				cg12513481		0.067	0.067	0.012	0.008	0.705
			HOXB2	cg25882366		0.260	0.275	0.044	0.046	0.640
			HOXB3	cg12910797		0.841	0.848	0.021	0.017	0.531
			HOXB4	cg02422694		0.040	0.040	0.004	0.004	0.907
				cg04609859		0.034	0.033	0.005	0.004	0.399
				cg06760035		0.033	0.034	0.003	0.003	0.031
				cg08089301		0.035	0.035	0.004	0.005	0.901
				cg14458834		0.118	0.127	0.019	0.023	0.082
				cg21460081		0.078	0.080	0.007	0.008	0.278
				cg21546671		0.115	0.120	0.016	0.020	0.607
				cg25145670		0.075	0.078	0.010	0.010	0.231
			HOXB5	cg01405107		0.273	0.234	0.041	0.054	2.11×10^{-5}
				cg16495265		0.081	0.081	0.008	0.009	0.861
			HOXB6	cg16848873		0.456	0.498	0.080	0.077	0.005
			HOXB7	cg06493080		0.051	0.050	0.008	0.009	0.912
				cg09357097		0.062	0.062	0.007	0.007	0.970
			HOXB9	cg12370791		0.030	0.031	0.003	0.003	0.559
				cg13643585		0.028	0.026	0.003	0.003	0.450
19	rs2363956	T/G	MYO9B	cg00215066	56/200	0.949	0.950	0.005	0.005	0.847
				cg01103836		0.962	0.963	0.003	0.003	0.869
			USE1	cg01252496		0.510	0.503	0.015	0.019	0.050
				cg08891270		0.035	0.034	0.002	0.003	0.882
			NR2F6	cg07242565		0.033	0.033	0.003	0.003	0.705
				cg16749578		0.166	0.164	0.022	0.023	0.901
			ANKRD41	cg21435336		0.037	0.035	0.003	0.003	0.004
			ABHD8	cg08145177		0.117	0.117	0.019	0.024	0.609
				cg22734480		0.027	0.025	0.002	0.003	0.038
			DDA1	cg19390148		0.032	0.031	0.005	0.005	0.534
				cg21061811		0.747	0.739	0.021	0.024	0.672
			MRPL34	cg04914305		0.050	0.050	0.005	0.004	0.538
	rs8170	T/C	MYO9B	cg00215066	180/76	0.950	0.950	0.005	0.005	0.515
				cg01103836		0.962	0.963	0.003	0.003	0.166
			USE1	cg01252496		0.505	0.505	0.018	0.017	0.598
				cg08891270		0.035	0.034	0.003	0.003	0.302
			NR2F6	cg07242565		0.033	0.034	0.003	0.003	0.646
				cg16749578		0.163	0.166	0.023	0.022	0.674
			ANKRD41	cg21435336		0.036	0.036	0.004	0.003	0.276
			ABHD8	cg08145177		0.118	0.113	0.022	0.026	0.320
				cg22734480		0.025	0.026	0.003	0.003	0.058
			DDA1	cg19390148		0.031	0.031	0.005	0.006	0.405
				cg21061811		0.735	0.749	0.027	0.019	0.403
			MRPL34	cg04914305		0.050	0.050	0.004	0.004	0.881

Table 5.13**:** Summary of genotype specific methylation of CpG islands associated with the extended list candidate genes within risk loci in chromosomes 17 and 19. Tabulated are the CpGs' median methylation (represented by β-values) of common

homozygotes compared to the median methylation of combined rare homozygotes and heterozygotes. P values were generated using the Wilcoxon rank sum test. Highlighted in yellow are P values that are significant before Bonferroni correction. Highlighted in green are statistically significant associations of methylation status and genotype after Bonferroni correction. CH: Common Homozygotes. RH+H: Rare Homozygotes + Heterozygotes combined.

Chromosome	Gene	CpG site	Linear regression analysis for age affecting methylation in healthy controls (P value)
2	HOXD9	cg10957151	0.334
		cg14991487	0.04
	HOXD8	cg15520279	0.039
		cg21815667	0.842
	HOXD3	cg00005847	0.295
		cg18702197	0.735
	HOXD1	cg19001226	0.047
	MTX2	cg19280014	0.684
3	KCNAB1	cg15423862	0.827
		cg23873703	0.786
	SSR3	cg24517609	0.776
	TIPARP	cg22114222	0.721
		cg24262469	0.665
8	MYC	cg19972619	0.094
		cg27207274	0.712
	PVT1	cg13052755	0.994
		cg13784855	0.008
9	BNC2	cg24341129	0.249
17	SP2	cg02691091	0.638
		cg14360917	0.487
	PNPO	cg00177698	0.578
		cg10154655	0.605
	NFE2L1	cg03777405	0.806
		cg06740897	0.893
	CBX1	cg11194725	0.268
		cg17778721	0.414
	SNX11	cg08413153	0.282
		cg12563480	0.985
	SKAP1	cg05106502	0.647
		cg12513481	0.02
	HOXB2	cg25882366	0.019
	HOXB3 HOXB4	cg12910797 cg02422694	0.408 0.782

		cg04609859	0.06
		cg06760035	0.013
		cg08089301	0.208
		cg14458834	0.677
		cg21460081	0.452
		cg21546671	0.744
		cg25145670	0.454
	HOXB5	cg01405107	0.148
		cg16495265	0.928
	HOXB6	cg16848873	0.173
	HOXB7	cg06493080	0.229
		cg09357097	0.068
	HOXB9	cg12370791	0.914
		cg13643585	0.391
19	MYO9B	cg00215066	0.051
		cg01103836	0.311
	USE1	cg01252496	0.64
		cg08891270	0.174
	NR2F6	cg07242565	0.46
		cg16749578	0.471
	ANKRD41	cg21435336	0.42
	ABHD8	cg08145177	0.134
		cg22734480	0.336
	DDA1	cg19390148	0.699
		cg21061811	0.515
	MRPL34	cg04914305	0.146
	MYO9B	cg00215066	0.051
		cg01103836	0.311
	USE1	cg01252496	0.64
		cg08891270	0.174
	NR2F6	cg07242565	0.46
		cg16749578	0.471
	ANKRD41	cg21435336	0.42
	ABHD8	cg08145177	0.134
		cg22734480	0.336
	DDA1	cg19390148	0.699
		cg21061811	0.515
	MRPL34	cg04914305	0.146

Table 5.14: Linear regression analysis of methylation in CpGs associated with the extended list of candidate genes' based on 3 age groups. Group 1: 50-59 yrs, Group 2: 60-69 yrs, Group 3: >70 yrs.

5.8 Discussion

5.8.1 Functional follow up of moderate risk susceptibility loci emerging from GWAS

Several GWAS have been performed over the last years that have identified multiple susceptibility loci for several common disease traits including several cancers. Although some risk variants lie in coding sequences, the vast majority reside in non-coding regions, including intergenic and intronic regions as well as gene deserts. Therefore, the observed associations cannot immediately provide us with a full understanding of the genetic mechanism underlying disease susceptibility, nor the genes involved. Post-GWAS functional characterisation studies in other types of cancer have identified genes that are implicated with disease development based on proximity to the risk loci. In breast cancer, after the identification of the risk locus 10q26, the gene *FGFR2* was implicated with disease susceptibility and later it was shown that the genotype of the risk variant located in the coding region of the *FGFR2* affects the transcript levels of the gene (Meyer et al, 2008). Several other genes have been implicated with cancer development after being investigated as candidates emerging form GWAS in breast, prostate, colorectal and lung cancer (Reviewed in Freedman et al, 2011).

So far the ovarian cancer GWAS performed has identified six confirmed low/moderate susceptibility loci for EOC. In the present study, I investigated whether these loci can provide insights into genes involved in EOC development. I have also looked at the mechanisms by which the risk associated SNPs may exert their susceptibility effect. These loci contain several genes that may be involved in EOC development. I investigated the functional relevance of candidate genes within these loci initially choosing the genes in the closest proximity with the EOC rick associated SNPs and then expanding the investigation to include more genes located up to 1Mb away of these variants. For this purpose I examined the differential expression of the selected candidate genes between a panel of NOSE and EOC cell lines. I also evaluated the functional significance of SNPs at these loci that are associated with EOC

susceptibility assessing whether their genotype is regulating *in-cis* the expression and methylation of candidate genes.

5.8.2 Functional role of candidate genes within the risk associated loci in EOC development

There is no reference framework in the research community on how to study the function of common low risk susceptibility loci, and so my initial aim was to establish the protocol and principles of these studies by investigating genes that were in the closest proximity to the most significantly risk associated SNP at each locus.

The first locus identified in the GWAS was locus 2q31 with the most significant SNP rs2072590, a non-coding SNP. This SNP is located within a non-coding hypothetical gene LOC401022 and there are no known functions for this SNP (http://PupaSuite.bioinfo.cipf.es/). This gene is found within the *HOXD* gene cluster between genes *HOXD1* and *HOXD3* (Goode et al, 2010). The *HOXD* gene cluster presents a family of transcription factors involved in morphogenesis, differentiation and development. Genes of this family have been previously shown to be implicated in lung adenocarcinomas, melanoma (Okubo et al, 2002), colorectal cancer and breast cancer and implicated more generally in neoplastic development with a loss of function role. More specifically, in colorectal cancer and lung adenocarcinomas a tumour suppressor role for *HOXD1* has been proposed as it was found to be epigenetically silenced through hypermethylation of CpG islands in this locus (Jacinto et al, 2007, Shiraishi et al, 2002). A genome wide screen recently showed that *HOXD1* is frequently hypermethylated and can serve as a reliable biomarker to detect early stages of breast cancer (Jeschke et al, 2012). It is only in an early study that gain of function role of *HOXD1* for neoplastic development has been proposed for human neuroblastoma by showing increased expression of HOXD1 in human neuroblastoma cell lines (Manohar et al, 1996).

The second locus identified was at 3q25 with the most significant SNP rs2665390 (Goode et al, 2010). There are no known functions for this SNP which is located within an intron of *TIPARP*. *TIPARP* is a gene coding for a

poly-ADP-ribose polymerase that may play a role in adaptive response to chemical exposure (Diani-Moore et al, 2010). Another member of the PARP superfamily is PARP1 which has a role in the DNA repair pathway. It has been reported that cells deficient in BRCA1 and/or BRCA2 can survive using the PARP pathway as an alternative DNA repair pathway (Farmer et al, 2005). PARP1 is also reported to be involved in transcriptional control and cellular differentiation (Masutani et al, 2002, Ji et al, 2010).

The third locus was 8q24 and the most significant risk associated SNP at this locus rs10088218. This SNP is an intergenic non-coding SNP with unknown function located downstream of *MYC* but in an apparent gene desert. This could be an indication that this SNPs is capable of regulating *MYC* from a distance or possibly that the real gene target regulated by the variant is not *MYC*. It is worth noting that 8q24 locus is rich in miRNA sequences that could be possibly the regulated targets for the variants of this locus. Other studies have also reported variants at 8q24 associated with susceptibility to several cancer types such as breast, prostate, colon and bladder cancers (Tenesa et al, 2008, Ghossaini et al, 2008, Kiemeney et al, 2008, Yeager et al, 2009, Al Olama et al, 2009). Most of the identified risk variants previously reported are located 5' of *MYC* but the EOC risk SNP lie 700kb 3' of *MYC*. However, I chose to investigate the *MYC* even though it wasn't the closest in proximity with the risk variant because it is a well-established oncogene and to evaluate whether this would be reflected on the expression data obtained.

The fourth locus was 9p22 with the most significant SNP rs3814113, an intergenic non-coding SNP with unknown function in close proximity to the gene Basonuclin-*2* (*BNC2*) (Song et al, 2009). A recent study found no evidence of an association of this SNP with ovarian cancer risk in women with endometriosis (Sundqvist et al, 2011) but another study has found an association with an increased risk of suspicious abnormal transvaginal ultrasounds in women with no ovarian cancer (Wentzensen et al, 2011). *BNC2* codes for a zinc finger DNA binding protein that is highly conserved and is suggested to be involved in DNA transcription. *BNC2* has been reported to have the potential to generate 90,000 different mRNA isoforms to be translated to 2,000 proteins (Vanhoutteghem et al, 2007). BNC1 and BNC2 have been

shown to be implicated in differentiation of oocytes and being highly expressed in ovary testis and skin keratinocytes, indicating a role in melanoma (Romano et al, 2004). The first report for *BNC2* related to cancer was that it has been shown in microarray analysis to be over-expressed in basal cell carcinoma compared to normal skin (O'Driscoll et al, 2006). BNC2 was proposed to act as a tumour suppressor gene after it was shown to be encompassed within a homozygously deleted locus and found under-expressed in glioblastoma, and causing growth arrest of oesophageal adenocarcinoma cells (Nord et al, 2009, Akaqi et al, 2009).

The fifth locus was at 17q21 with the most significant SNP rs9303542. There are no known functions for this SNP and it is located within an intron of *SKAP1*. SKAP1 has been shown to regulate mitotic progression at the transition from metaphase to anaphase (Fang et al, 2009). Low expression of SKAP1 has been implicated with poor prognosis in breast cancer and SKAP1 function is proposed to predispose to sensitivity to paclitaxel by maintaining chromosome stability and mitotic fidelity (Burell et al, 2009). SKAP1 has recently been shown to be involved in regulating the NF-κB pathway during T-cell activation. NF-κB activates the expression of genes that keep the cell proliferating (Burbach et al, 2011).

The sixth locus was 19p13 with rs8170 and rs2363956 identified initially in association with survival in ovarian cancer case but later on also shown to be associated with disease risk driven by the serous subtypes (Bolton et al, 2010). rs8170 and rs2363956 are non-synonymous coding SNPs with unknown function located within *MERIT40* and *ANKRD41* genes respectively. No known function for *ANKRD41* is reported in the literature. *MERIT40* is a very interesting gene as its coded protein has been reported to interact with BRCA1, RAP80, BRCC45 and CCDC98. Through this interaction MERIT40 regulates the recruitment of BRCA1 at double strand DNA breaks, maintains stability of this complex at the sites of DNA damage and has been shown to play a role in cell cycle progression (Feng et al, 2009, Shao et al, 2009, Wang et al, 2009).

In order to evaluate the functional role of the candidate genes at these loci in EOC development I compared the expression of each candidate gene in 48 NOSE versus 24 EOC cell lines. I found that there was significantly higher

expression of *BNC2* and *TIPARP* in NOSE cell lines compared to EOC cell lines which may indicate a tumour suppressor role for these two genes. On the other hand, the expression of *HOXD1*, *MYC*, *SKAP1* and *MERIT40* was significantly higher in EOC cell lines compared to NOSE cell lines perhaps suggesting a gain of function role for these genes in EOC development, something that is already well-established for *MYC*. *HOXD1* is reported to act as a tumour suppressor gene in several other cancers and my study is the first one supporting an a gain of function role of this gene in carcinogenesis.

Expression microarray analysis of an *in vitro* model of neoplastic transformation of IOSE cells supported some of the findings from my expression analyses; *BNC2* and *TIPARP* expression decreased significantly with each additional oncogenic event, as the cells acquired a more neoplastic phenotype. The proposed loss of function role for those two genes was also strongly supported by showing higher expression in normal FT tissue compared to serous ovarian carcinomas. The observed over-expression of *SKAP1* in EOC cell lines was not consistent with the *in vitro* IOSE transformation model, where it was not statistically significant differentially expressed, although there was a trend towards increased expression in progressively transformed IOSE cells. By contrast, *SKAP1* appeared to be highly expressed in the normal FT from TCGA data compared to EOC suggesting instead a tumour suppressor role for this gene in EOC development. Additionally, the observed over-expression of *MYC* and *HOXD1* in EOC cell lines was not consistent with the TCGA data where *MYC* and *HOXD1* did not appear to be differentially expressed between normal FT and SOC. Because *MYC* is a well-established oncogene in EOC development; while there was marginal, significant over-expression of the gene in EOC cell lines, one might have expected this to be more significant. With *HOXD1*, the observed over-expression in EOC should be treated with more caution: Firstly, the expression of the gene in both EOC and NOSE cell lines was just marginally detectable; secondly, several previous reports in other types of cancer for this gene suggest it acts as a tumour suppressor gene.

The contradictions between the different models might be attributed to the nature of the models analysed or simply chance observations. The IOSE *in vitro* transformation model does not resemble serous events and that might be

the explanation for the differences observed since most of the EOC cell lines I used were of serous subtype. As previously discussed, the normal FT mRNA used by TCGA was obtained from whole fallopian tube tissue rather than from the specific cells of origin for EOC, which are the fallopian tube epithelial cells probably of secretory lineage. Thus it is conceivable the expression of many genes in the epithelial and stromal cells contained in these samples may considerably vary and due to the high proportion of stromal cells to provide misleading results when compared with SOC samples. Additionally, the EOC cell lines used were of several subtypes and may not represent an ideal model to study a gene that is involved in development of certain EOC subtypes only. Finally, the expression data from TCGA were generated using gene expression microarrays and it may be that experimental variation between RT-PCR and array approaches can produce different results. Previous research has shown that the differential expression of genes from microarray analysis is not always in agreement with Real time PCR results (Kothapalli et al, 2002).

Finally, *MERIT40* over-expression observed in EOC cell lines was not consistent with the in *vitro* transformation model of IOSE cells. However, this could just mean that MERIT40 and HOXD1 are involved in EOC development through another pathway than the MYC/KRAS pathway that was used to transform the IOSE cells. *MERIT40* was also significantly over-expressed in SOC supporting the gain of function role for this gene I propose based on the EOC versus NOSE expression model. Additionally, supporting the gain of function role for *MERIT40*, Dr Chris Jones within our group has performed array Comparative Genomic Hybridization (aCGH) analysis to evaluate genomic alterations at the 19p13 locus containing *MERIT40*. This analysis revealed copy number gain/amplification of the p-arm of chromosome 19 in 45% of 105 tumour samples. This may suggest that target genes in this region including *MERIT40* are functionally activated during tumour development. Additionally, according to TCGA, SOC samples that were found to have high expression of *MERIT40* are shown to have copy number gains and amplifications at this locus (www.cbioportal.org).

Initially, I chose candidate genes for analysis that were in close proximity to the risk variants within the six loci. Based on the results previously described

TIPARP, HOXD1, MYC, BNC2, SKAP1, MERIT40 were proposed to be involved in EOC development. However, it is possible that other genes within these loci have a higher or an additional significance in the development of EOC. Therefore, I performed a multiplex Fluidigm gene expression assay investigating several genes in each locus including the ones investigated in the initial study in the closest proximity to the risk variants. This also allowed me to validate the results of my initial analysis of genes in closest proximity in a larger panel of cell lines. I also added fallopian tube epithelial (FTE) cell lines (n=5) to the panel given the proposed dual origin for ovarian cancer to comprise the normal cell lines that their expression would be compared to a larger panel of 44 EOC cell lines.

Of the 62 genes investigated in total 33 only remained after quality control analysis. One should not be too concerned regarding the failure of almost 50% of the assays in this experiment because it is possible that many of the genes investigated did not have a high enough expression in ovarian cell lines to be detected as gene expression is tissue specific.

After analysing the remaining additional candidate genes in locus 2q31, *MTX2* (Metaxin 2), a gene located 1kb downstream rs2072590, seems to be an interesting candidate as it was found over-expressed in EOC cell lines compared to normal cell lines consistent with the TCGA expression analysis. The result did not remain significant after Bonferroni correction though so should be treated cautiously. *MTX2* is coding for a protein involved into transporting proteins into the mitochondrion. MTX2 abnormal expression has been previously reported to be associated with poor prognosis linked to resistance in a distinct group of acute myeloid leukaemia (AML) patients that have normal cytogenetics and was difficult to identify with conventional cytogenetic analysis (Vey et al, 2004).

The investigation of additional genes in locus 3q25 revealed that *TIPARP* remains the most significantly associated gene with EOC development within this locus being very significantly under-expressed in EOC compared to normal cell lines ($P= 4.99\times10^{-5}$) and also in SOC compared to normal FT ($P= 1.11\times10^{-4}$). Additionally, the gene *PA2G4P4* (proliferation associated 2G4 pseudogene) was found significantly over-expressed in EOC cell lines compared to normal

cell lines (P= 2.14×10^{-5}). This gene has a great homology with the mouse cell cycle-protein p38-2G4 involved in proliferation (Lamartine et al, 1997). There is no established function reported for this gene or any known implications in cancer.

Interestingly, for locus 8q24 the previously observed slight over-expression of *MYC* in EOC cell lines did not remain significant in the extended analysis although a similar trend was observed in the boxplot (Figure 5.13 B). The other gene studied on this locus, *PVT1* (plasmacytoma variant translocation 1) is a gene downstream *MYC* in closer proximity to the risk associated SNP and was found to be the most functionally significant gene in this region with a very significant over-expression of the gene observed in EOC compared to the normal cell lines (P= 6.6×10^{-8}).

The exact functional role of *PVT1* has been enigmatic for more than two decades since the gene was identified because although *PVT1* gene directs the synthesis of a large transcript no protein product had been identified. A cluster of putative micro RNAs (miRNAs) (miR-1204~1208) has been identified within the *PVT1* genomic locus with a functional role in T lymphomagenesis in Burkitt's lymphomas (Huppi et al, 2008, Beck-Engeser et al, 2008). This gene is reported to act as a potential oncogene. It was found to be over-expressed in several transformed cell lines such as neuroblastoma cells (Carramusa et al 2007). The inactivation of *PVT1* has also been shown to restore sensitivity in Gemcitabine chemotherapeutic drug in pancreatic cancer (You et al, 2011). It has also been shown to from a fusion gene with *CDH7* (cadherin 7), coding for a membrane protein, which is found amplified in small cell lung cancer (Pleasance et al, 2010).

According to TCGA 30-32% of a total of 316 SOC had amplifications in the *MYC* and *PVT1* loci (www.cbioportal.org). The observed over-expression for *PVT1* but not for *MYC* in the EOC cell lines was puzzling. However, despite the long history of *MYC* as an oncogene there is a lot of recent evidence that shows that *PVT1* may be the driver gene of the amplifications found in this locus in EOC. One study has shown that *PVT1* but not *MYC* is strongly over-expressed in ovarian tumours with genomic amplifications compared to normal FT (Haverty et al, 2009). It has also been shown that knocking down *MYC* and *PVT1* in

ovarian cancer cell lines caused reduced proliferation but only *PVT1* knockdown caused increased apoptosis. In the same study it was shown that some EOC cell lines did not overexpress *MYC* whereas transcription levels of *PVT1* were high in all cell lines, and that in EOC cell lines with no amplification of the 8q24 locus, *MYC* was not over-expressed but *PVT1* was (Guan et al, 2007). Thus, it is possible that in the increased number of EOC cell lines used in the extended analysis I performed, more cell lines were incorporated with low *MYC* expression and masked the over-expression previously observed in the smaller subset.

The induction of *PVT1* transcript synthesis is not clear. Guan and colleagues have shown that *PVT1* and *MYC* act independently (Guan et al, 2007) but other studies have shown that MYC binds to the promoter of *PVT1* activating its transcription (Carramusa et al, 2007) whereas others support that *PVT1* encoded miRNAs are regulating *MYC* itself (Huppi et al, 2008). A recent study has shown that *PVT1* is a p53-inducible target gene (Barsotti et al, 2012). A recent review has emphasized the significance of locus 8q24 in carcinogenesis as being gene desert rich in non-coding regulatory elements such as miRNAs and raises awareness of other transcripts besides *MYC* in the 8q24 region as possible candidate genes in malignancy formation (Huppi et al, 2012). My finding is supported by the relevant research performed for *PVT1* proposing an important role in ovarian cancer development and in other malignancies.

In locus 9p22, *BNC2* remained the most significant candidate gene as I found it was over-expressed in EOC cell lines compared to normal ($P= 6.64\times10^{-4}$) confirming the loss of function role proposed for this gene in the pilot study.

Locus 17q21 is a gene rich region and several genes were selected to be investigated in a region of ~1.2 Mb. The data for the candidate genes of this locus were often inconsistent between the EOC versus normal analysis and the SOC versus normal FT analysis from TCGA. Three genes of the HOXB (homeobox B) cluster of genes were found to be significantly over-expressed in EOC cell lines compared to normal but in contrast appeared under-expressed in SOC compared to normal FT according to the TCGA expression data analysed. The genes with the observed contradicting expression were *HOXB3*, *HOXB4*

and *HOXB6*, the results for which remained significant after Bonferroni correction for both models studied. *HOXB7* and *HOXB9* also were found over-expressed in EOC cell lines but no differential expression was found in the SOC compared to normal FT.

HOXB genes function as transcription factors for the developmental regulatory system. *HOXB3* and *HOXB4* have been previously found to be expressed significantly higher in ovarian cancer cell lines compared to normal ovaries based on mRNA (Yamashita et al, 2006) and protein expression analyses (Hong et al, 2010) in line with the results of this study. *HOXB4* has been also found to be over-expressed in cervical cancer compared to normal adult cervical epithelium (Lopez et al, 2006). *HOXB3* has been shown to have abundant expression in breast cancer and regulated by miRNAs plays a critical role in the epigenetic silencing of tumour suppressor genes inducing the growth of transformed cell lines (Li et al, 2012) Additionally, *HOXB3* expression has been found elevated in acute myeloid leukaemia (AML) (reviewed by Eklund, 2011) and involved in chemoresistance of AML cases (Kuhnl et al, 2011). There are no previous reports for a role of *HOXB9* in ovarian cancer. Elevated expression of HOXB6 was found to be a common event in ovarian tumours after serologic analysis of ovarian tumour antigens (Stone et al, 2003). HOXB7 over-expression has been shown to increase proliferation of IOSE cell lines and has been found expressed in markedly higher levels in ovarian carcinomas compared to normal ovarian surface epithelium (Naora et al, 2001). Additionally, knocking down *HOXB7* in ovarian cancer cell lines has reported to cause an 85% reduction to their invasion abilities (Yamashita et al, 2006). All the previous work proposes an involvement in malignancy formation for all the studied *HOXB* genes in several types of cancers included ovarian cancer suggesting that they may act as oncogenes and this is in line with the gain of function role in EOC proposed for the *HOXB* genes in this study.

Three more genes of this locus were found to be over-expressed in EOC compared to normal cell lines, suggesting a possible gain of function role in EOC for them, *PNPO, SP2* and *CBX1*. *PNPO* (pyridoxamine 5'-phosphate oxidase) is coding for a protein that is involved in the formation of the active form of Vitamin B6 which has been shown to confer sensitivity of breast cancer

cells to tamoxifen (Aldhaheri et al, 2006). The functional role of PNPO in malignancy formation has not been studied. *SP2* is coding for an evolutionary conserved transcription factor for genes that are involved in development and cell cycle progression and it has been shown that it was required for normal progression of embryonic development in mice (Baur et al, 2010). To my knowledge there are no extensive evidence suggesting the implication of SP2 with human cancer development. There is one report suggesting that the expression of SP2 is directly correlated with the progression of murine squamous cell carcinomas. In this study, the observed SP2 over-expression inhibited keratinocyte differentiation rendering those cells susceptible to neoplastic development proposing a role for *SP2* as an oncogene (Kim et al, 2010). *CBX1* (chromobox homolog 1) is coding for one of the heterochromatin proteins that are involved in chromatin packing and gene regulation. An important finding regarding CBX1 has been the report that following DNA breaks it is mobilized and released from chromatin to initiate DNA damage response (Ayoub et al, 2008). Thus, it is attractive to speculate that the *CBX1* over-expression in EOC cell lines may suggest that the gene involved in EOC through activation of DNA damage response pathways inducing DNA repair in late stages of EOC rendering cancer cells resistant to platinum agents.

MERIT40 in locus 19p13 was considered a strong candidate gene for association with EOC development due to its proposed function in homologous recombination as part of the BRCA1 complex and as it was found to be over-expressed both in the EOC cell lines versus NOSE in the pilot and extended studies consistent with over-expression in SOC versus normal FT. The over-expression of *MERIT40* in EOC cell lines may seem as an apparent paradox to be linked with EOC susceptibility since BRCA1 is expected to show loss of function in its role in the repair or DNA DSBs for EOC initiation. One hypothesis would be that over-expression of *MERIT40* may ectopically stabilize mutant BRCA1 protein into the assembled complex. Since *MERIT40* makes cells more resistant to ionizing radiation (Shao et al, 2009), *MERIT40* over-expression may be implicated in EOC initiation by protecting cells with dysfunctional BRCA1 and DNA DSB repair activity and enabling them to tolerate more DNA damage. However, it is probable that the expression analysis performed could be

unrelated to the susceptibility SNPs and could just be an indication of the somatic role of the candidate genes in EOC development rather than susceptibility. Thus, the possible functional role of MERIT40 in the initiation of serous subtype EOCs and if it is the target susceptibility gene at the 19p13 locus can only be speculated. Since these results are probably reflecting the somatic role of *MERIT40* in EOC development the proposed gain of function role for this gene in EOC may be exerted by providing cancer cells increased ability to repair DNA DSBs caused by spontaneous endogenous stress or platinum drugs thus promoting progression and survival in EOC tumours. This hypothesis is also supported with previous reports that have shown BRCA1 to be reactivated in advanced tumours providing chemoresistance to platinum therapy. Future work to elucidate the possible functional role of *MERIT40* in EOC development seemed very attractive, and this has been pursued in chapter 6 of this thesis.

The most statistically significant difference in expression for this locus, were for the genes *HAUS8* and *USE1*. Both were found to be over-expressed in EOC compared to normal cell lines indicating a gain of function role for those genes in EOC development ($P= 2.09E\times10^{-5}$ and $P= 1.08E\times10^{-6}$ respectively). *HAUS8* was also found over-expressed in SOC compared to normal FT but *USE1* was not differentially expressed between them. HAUS8 is a centrosomal microtubule-binding protein involved in mitotic spindle assembly and the process of cytokinesis (Wu et al, 2009). *USE1* function is not very understood yet but it may function in the Golgi apparatus and in endosome lysosome transport (www.genecards.org). To my knowledge there are no reported implications of *HAUS8* or *USE1* genes with cancer.

5.8.3 Evaluating genotype specific gene expression and genotype specific methylation for the GWAS candidate genes

In silico analyses in PupaSuite of risk-associated and strongly-correlated SNPs (r^2>0.8) failed to find compelling evidence for any functional role. I therefore investigated the possible functionality of the SNPs that were found in association with EOC susceptibility into regulating *in cis* the transcriptional output of candidate genes in each respective locus. For this, I evaluated whether there was any genotype specific expression of the candidate genes relative to the genotypes of the significant SNPs.

Increased DNA methylation and reduced tumour expression implicated more than 100 of genes as epigenetically silenced in high grade serous ovarian cancer samples compared with fallopian tube controls according to TCGA. Increased DNA methylation has been shown to be correlated with reduced gene expression across all the samples (The Cancer Genome Atlas Research Network, 2011). Thus, based on the critical role methylation status plays in gene expression, I investigated whether the genotype of the risk associated SNPs affects the expression of the candidate genes and whether there would be a link between genotype affecting gene expression and methylation of the same genes.

Initially, in a pilot study I investigated the expression of the seven candidate genes I selected on close proximity to the risk associated SNPs relative to the genotypes of those SNPs in 35 NOSE cell lines. I found no evidence of genotype specific expression for the genes initially analysed. I then decided to investigate genotype specific gene expression for the close proximity candidate genes in a greater number of genotyped samples for the risk associated SNPs (n=56 NOSE and FT samples). Additionally, I extended the study and analysed genotype specific gene expression of the whole panel of selected candidate genes across the six loci. I did not find any significant associations in the pilot study but I found a few in the extended study which however did not remain significant after Bonferroni correction for multiple testing. Regardless, the results will be discussed assuming that increasing the number of tested samples in the future may improve the power of the study.

The first genotype association with gene expression I found was between expression of *HOXD1* and the genotype of the rs2072590 in locus 2q31. Individuals carrying the minor allele T (rare homozygotes TT and heterozygotes GT) had over-expressed *HOXD1* (P= 0.041) compared to individuals carrying the common genotype (GG). It is interesting that *HOXD1* was found to be over-expressed in EOC cell lines compared to normal proposing a gain of function role in EOC development. This observation together with the fact that the minor allele was associated with an increase in the transcript output of *HOXD1* may suggest that the reported in the GWAS increased risk for susceptibility in EOC of the minor allele of rs2072590 might be associated with a functional role of this SNP by regulating the expression of a potential oncogene, *HOXD1*. The result should though be validated in the future in a larger sample size.

Genotype specific methylation analysis was performed to reveal whether the mechanism of a potential functional role of rs2072590 is exerted by epigenetically regulating the expression of the gene. In the pilot study performed to investigate genotype specific methylation I evaluated genotype specific methylation of CpGs of 148 healthy individuals in the candidate genes in close proximity to the significant SNPs including rs2072590. This analysis produced a marginally significant result proposing genotype specific methylation for cg19001226 of *HOXD1* (P= 0.035) for SNP rs2072590. Individuals carrying the common (GG) genotype were found to be hypomethylated in cg19001226 proposing that the minor allele was associated with hypermethylation of cg19001226 which may lead to epigenetically decreased expression of the *HOXD1*. This would pose as a contradiction since the minor allele is the risk associated allele and *HOXD1* was found over-expressed in EOC cell lines posing as a potential oncogene for EOC. However, in the extended study I revaluated the methylation status of cg19001226 in 256 healthy individuals the rs2072590 genotype association with methylation was not validated which is why I conclude that the regulatory role of rs2072590 is probably not exerted by epigenetically regulating the expression of the gene.

SNP rs2072590 is located within the non-protein coding RNA gene LOC401022 between *HOXD1* and *HOXD3*. This non-coding gene is also called HOXD cluster antisense RNA and is a long antisense non-coding RNA which is

a non-coding transcript longer than 200 nucleotides and that is what is differentiating this group of non-coding RNAs from small regulatory RNAs such as miRNAs. There is no reported function for this non-coding RNA (http://lncrnadb.com). Long non-coding RNAs exist in abundance in the human genome and may serve as primary transcripts for the production of short RNAs (Kapranov et al, 2007). Some have been reported to have a role in the regulation of gene transcription by modulating the function of transcription factors. For example, the long non-coding RNA Evf-2 is functions as a co-activator with the homeobox transcription factor Dlx2 (Feng et al, 2006).

Thinking of the potential mechanism by which rs2072590 exerts its regulation on *HOXD1* expression one could speculate that it may be through regulating the production of LOC401022 which in turn could act as transcriptional regulator of *HOXD1.* The common allele might be responsible for producing the antisense transcript that causes post transcriptional regulation of both copies of HOXD1 via RNA interference and thus individuals carrying the minor/risk allele might deregulate the *HOXD1* post-transcriptional inhibition which could partially explain the increased risk to EOC susceptibility if we accept the proposed role of *HOXD1* as an oncogene in EOC. This hypothesis is supported by a recent study that has identified associations of miRNAs expression and the genotype of rs2072590 (Shan et al, 2012). Following this report I have investigated which miRNAs were found to be regulated by rs2072590 in this study and have looked in publicly available miRNA function databases to try and identify whether any of those miRNAs would be reported to target *HOXD1*. I did not identify exact matches but miRNAlet-7a* (let-7 miRNA produced from the 3' of the transcript) was differentially expressed relative to rs2072590 genotype ($P= 5\times10^{-4}$) (Shen et al, 2012) and miRNAlet-7a (let-71 miRNA produced from the 5' of the transcript) is reported to target *HOXD1* (http://mirdb.org/miRDB). Increasing research is supporting the involvement of non-coding RNA transcripts in human tumorigenesis and non-coding RNAs have been found with altered expression in several types of cancer such as uterine cancer, hepatocellular carcinomas, leukaemia and ovarian cancer (Calin et al, 2007, Mallardo et al, 2008).

The most interesting genotype specific gene expression result, though not significant after Bonferroni correction, was the effect of rs9303542 genotype on the expression of three genes in locus 17q21 *CBX1, SNX11* and *SP2*. Individuals with the minor allele (G) that was associated with an increased risk for EOC susceptibility in the GWAS, were found to have *CBX1* over-expressed (P= 0.006). This gene was found over-expressed in EOC cell lines compared to normal cell lines and the minor allele might be conferring the increased risk to EOC partially by elevating the expression of the potential oncogene *CBX1*. Given the role of CBX1 in triggering DNA damage responses the genotype specific gene expression would suggest that rs9303542 may contribute to interindividual variability in drug response. Identification of such variants could be valuable as they could be used as biomarkers for individualising anticancer therapy. However, the latter hypothesis could be under debate since rs9303542 has not been found associated with survival but only susceptibility in EOC.

No association of rs9303542 genotype was found with methylation in the *CBX1* gene locus. However, genotype specific methylation analysis showed that rs9303542 genotype was associated with methylation of cg0145107 of *HOXB5*. The minor allele was significantly associated with hypomethylation of cg0145107 ($P= 2.11\times10^{-5}$), which could suggest it may also be associated with increased expression of *HOXB5*, assuming there is a correlation between expression of *HOXB5* and methylation for this CpG island. HOXB5 functions as a transcription factor for the developmental regulatory system. HOXB5 was found to be implicated with promoting differentiation in alveolar epithelial cell of the lung (Fu et al, 2008) and involved in the differentiation of the vascular endothelium from precursor cells (Wu et al, 2003). The expression assay for *HOXB5* failed QC in this study, or the expression is very low in ovarian tissue to be detected, and thus I cannot propose a role of the gene in EOC development based on differential expression between the EOC and normal cell lines. *HOXB5* was found under-expressed in SOC (P= 0.002) supporting a loss of function role for the gene in EOC development. Additionally, *HOXB5* was found to be hypermethylated in ovarian carcinomas also suggesting that it may act as a tumour suppressor gene (Wu et al, 2007). However, there is a recent report though that supports *HOXB5* may be acting as an oncogene as it was found

over-expressed in oral squamous cell carcinoma (Tucci et al, 2011). Based on other reports and results of this thesis suggesting a potential tumour suppressor role of *HOXB5* in ovarian cancer it seems a paradox that the risk associated minor allele of rs9303542 would be associated with epigenetically activating the expression of the gene. However, there is no evidence that this CpG is definitely affecting expression levels of the gene and no association of genotype and gene expression was found in this study. The biological basis of ovarian cancer development is very complicated and more work needs to be performed to elucidate the role of *HOXB5* in ovarian carcinogenesis and the variants that may be regulating its function before drawing any conclusions.

The rs9303542 minor allele was also found to be associated with the decreased expression of *SNX11* and *SP2* (P= 0.028 and P= 0.03 respectively). *SNX11* and *SP2* were significantly over-expressed in EOC cell lines (P=0.004 and P= 5.6×10^{-10} respectively) suggesting a gain of function role for the genes. Again, it is puzzling why the allele that is conferring increased risk in ovarian cancer susceptibility would be associated with decreased expression of candidate oncogenes. It could be possible that allele specific LOH targeting the minor allele could be one mechanism involved in the increased expression of SP2 observed in EOC cell lines but that should be further investigated. Regardless of the difficulties in fully understanding the functional role of rs9303542 in the development of EOC it is very intriguing that the genotype of this SNP is affecting the expression of three genes and the methylation of another.

Genotype specific methylation has not been extensively studied. There is one study that in the attempt to show that DNA methylation is subject to genetic control, reports very significant associations of SNPs' genotype with DNA methylation levels near the polymorphism (*cis*-effects) and also weaker associations with DNA methylation in a further locus (*trans*-effects) (Boks et al, 2009). Another study has used HapMap derived cell lines and identified associations between genotype and DNA methylation and gene expression at a genome-wide level. Their most significant finding was that the genotype of rs2187102 was associated with methylation status in a CpG in gene *HLCS* which is involved in gene regulation by mediating histone biotinylation. The

most interesting point of this study was that they observed a significant overlap between regulatory variation that affects both methylation and gene expression. Their hypothesis was that since DNA methylation can regulate gene expression then SNPs that are regulating methylation often would have effects seen on gene expression for the same locus (Bell et al, 2011). I have made a similar observation regarding *HOXD1* methylation and expression being associated with the genotype of rs2072590 but unfortunately the methylation association did not remain significant when I increased the power of the study. In order to test the hypothesis regarding correlation of genotype specific methylation and gene expression, I would be very interested to know the potential genotype specific gene expression for *HOXB5* for the SNP that was found to affect so significantly its methylation but the gene assay failed in the experiment performed.

Several previous studies have also shown evidence of genotype specific expression in an attempt to elucidate the possible functional role of genetic variants. Keratin1 (*KRT1*) has been found to have extremely significant genotype specific expression in human white blood cells by SNPs that had differential affinity for transcription factors modulating *KRT1* promoter (Yao et al, 2006). Another study has shown differential expression of several genes in human white blood cells regulated by *cis*-elements in LD with the SNP of interest and proposed allele specific expression of these genes (Pant et al, 2006). A recent study has showed that the risk allele of rs378854 is associated with increased expression of the *PVT1* oncogene in locus 8q24. In this study they used mapping of DNase I hypersensitive sites in order to prioritise the regions that would be interesting to further analyse for their functional role (Meyer et al, 2011). In a study for breast cancer, microarray analysis was used to identify that the expression of *FGFR2* gene is increased in the rare homozygote samples. They proposed that two *cis*-regulatory SNPs alter the affinity of transcription factors and synergistically they cause increase expression of *FGFR2* (Meyer et al, 2008).

Interestingly, the functional role of the risk associated SNPs that the ovarian cancer GWAS identified has been recently studied in regards with their associations with miRNAs in ovarian cancer. miRNAs act as post transcriptional

regulators by targeting mRNAs and causing gene silencing. Thus, they can function as oncogenes or tumour suppressor genes and they have recently been implicated with ovarian cancer development. A recent study has studied the effect of the genotype of the ovarian cancer GWAS' identified most significant SNPs in 2q31, 3q25, 8q24, 9p22 and 19p13 and reported significant association of miRNA expression and the genotype of the EOC risk associated SNPs rs3814113 and rs2072590 in loci 9p22 and 2q31 respectively (Shen et al, 2012). This study offered a new insight into the functional role for risk variants emerging from GWAS and they are the first to report that EOC risk assciated SNPs regulate miRNas. The associations they found between miRNA expression and SNP genotype in this study did not always correlate positively with the biological functions of the miRNAs, which is something that I also found for several of the functional associations I observed in my study.

5.8.4 Conclusion

In conclusion, functional analysis of the candidate genes in the EOC risk associated loci has provided with information on the genes potential involvement in EOC. I found that *PVT1*, in locus 8q24, *SP2, CBX1*, *PNPO* and *SKAP1* and several genes of the HOXB family in locus 17q21, *HAUS8, USE1* and *MERIT40* in locus 19p31 are over-expressed in EOC cell lines compared to normal NOSE and FTE cell lines suggesting a gain of function role in EOC. Additionally, I found *TIPAPRP* in locus 3q25, *BNC2* in locus 9p22 to be under-expressed in EOC cell lines suggesting a loss of function role for these genes in EOC development. Some of these genes such as *TIPARP*, *PVT1*, and *MERIT40* may have an intriguing role in EOC based on previous reports regarding their biological function and should be investigated in more gene specific functional assays to further confirm their potential role in EOC development.

This study reports, for the first time, genotype specific gene expression and methylation of genes within risk associated loci in ovarian cancer. To summarise, weak associations were found between the genotype of rs9303542 and the expression of *CBX1*, *SNX11* and *SP2* genes in locus 17q21 and

rs2072590 and *HOXD1* expression in locus 2q31. Additionally, methylation of *HOXB5* was associated with the genotype of rs9303542 in locus 17q21. These results suggest that the genes regulated by the genotype of risk associated SNPs may be the target susceptibility genes for the relevant loci.

An advantage of the study lies in the fact the genotype specific gene expression was performed using a proposed cell of origin for EOC, which in turn is the limitation for the genotype specific methylation analysis where germline DNA was used since expression and methylation are tissue specific. Another limitation of the study is the number of the samples used may not have been large enough to generate the power to identify association. Another limitation of this study regarding the investigation of functional relevance of the risk associated SNPs in EOC development is that they are not for sure the true causative SNP and another SNP in LD with the associated SNP may be the true functional variant. Several questions still pose, such what are the exact mechanisms that SNPs are exploiting to be functionally involved with EOC development be regulating the expression and methylation of candidate genes. Further understanding of the molecular basis of the SNPs mechanisms to affect the epigenetic and transcriptional regulation of genes would shed some light into the molecular basis of risk etiology.

♣ Publications containing work from this chapter:

Goode EL*, Chenevix-Trench G*, Song H*, Ramus SJ*, **Notaridou M**, Lawrenson K, Widschwendter M et al, "*A genome-wide association study identifies susceptibility loci for ovarian cancer at 2q31 and 8q24*", Nat Genet. 2010 Oct;42(10):874-9.

Bolton KL*, Tyrer J*, Song H, Ramus SJ, **Notaridou M**, Jones C, Sher T et al "Common variants at 19p13 are associated with susceptibility to ovarian cancer", Nat Genet. 2010 Oct;42(10):880-4.

* Authors contributed equally to this work

6 Investigation of the potential role of *MERIT40* in EOC development

6.1 Introduction

Following the identification of the risk associated loci from the GWAS I have proposed a functional role for several candidate genes within these loci in EOC development. I have chosen to take *MERIT40* (recently renamed as *BABAM1*) forward in order to try and elucidate by phenotypic *in vitro* assays its specific potential biological role in EOC. The reason I selected this gene, and not some other that was more statistically significant in the differential expression analysis between the EOC and normal cell lines, was the intriguing role it was recently proposed to have in DNA repair as part of the BRCA1-Rap80 complex. MERIT40 has been shown to facilitate the localisation of the complex in DNA double strand breaks (DSBs) to initiate DNA repair by homologous recombination (HR) (Feng et al, 2009, Shao et al 2009). In the event of DNA DSBs the cell responds by recruiting various protein complexes to the site of the damage in order for HR to repair of the lesion. The tumour suppressor protein BRCA1 has an important role in the HR pathway and is recruited to the site of a DNA DSB as part of distinct protein complexes involved both in DNA DSB recognition and resection. BRCA1 binds to DNA DSBs through its association with the Abraxas(CCDC98)-Rap80-BRCC36 complex which associates with ubiquitinated histones at DNA DSBs following phosphorylation of histone γH2AX. Recent studies have reported that MERIT40 interacts with Rap80 through its CCDC98 interacting region and in order to facilitate the BRCA1-Rap80-CCDC98 complex integrity and DNA DSB localisation. These studies showed that MERIT40 is essential for the G2 checkpoint and for cellular resistance to ionising radiation through facilitating the recruitment of BRCA1 to the DNA damage foci (Shao et al, 2009, Feng et al, 2009). The expression of MERIT40 has been interlinked with maintaining the levels of components of the BRCA1-CCDC98-Rap80-RCC36 complex (Hu et al, 2011).

Homologous recombination is a DDR pathway commonly deregulated in epithelial ovarian tumours with inactivating *BRCA1* mutations. Tumours with such mutations respond well to platinum therapy such as cisplatin and carboplatin (Agarwal and Kaye, 2003). Cisplatin induced DNA damage is activating a number of pathways that lead to apoptosis and death of the cancer cell. The cytotoxicity of cisplatin is attributed to its ability to interact with purine bases on the DNA backbone leading to the formation of DNA-protein and DNA-DNA interstrand and intrastrand crosslinks. The interstrand DNA adducts formed are initiating a sequence of steps leading to the completion of the cytotoxic process, namely apoptosis (Pinto and Lippard, 1985). The DNA damage is recognised by proteins that bind to physical distortions in the DNA induced by the platinum adducts and transduce DNA damage signals to downstream effectors to prevent transcription and replication and lead to apoptosis (McCabe et al, 2009, reviewed in Siddik, 2003). The main DDR pathways involved in the repair of platinum adducts are the NER and HR pathways. This repair is a multistep process where firstly an endonucleolytic excision takes place in one strand, followed by translesion synthesis across the interstrand crosslink and removal by excision. The resulting DNA DSBs should then be resected to generate ssDNAs followed by repair by HR. BRCA1 has been shown to be involved in transcription coupled NER (Gowen et al, 1998) and has a central role in HR. Thus in BRCA1 deficient cells these mechanisms are defective rendering the cells hypersensitive to platinum agents.

Although defective DNA repair causes for an initial good response after treatment of EOC with platinum agents, restoration of DNA repair results to an acquired resistance to these drugs. Previous work has shown that secondary activating mutations on *BRCA1* and *BRCA2* are found in ovarian carcinomas with platinum resistance (Swisher et al, 2008).

Based on the role of MERIT40 being a BRCA1 associated protein involved in DNA repair, I hypothesised that potentially the gain in function role I proposed in EOC development for MERIT40 may be exerted by providing the cancer cells with tolerance against DNA damage and increased resistance to platinum drugs. I also investigated whether *MERIT40* would have another

biological relevance to the development of EOC as an oncogene by regulating cell cycle regulation, proliferation or migration.

Aims of this chapter:

1. Create EOC cell lines with stably silenced *MERIT40*
2. Investigate whether silencing *MERIT40* in EOC cell lines will be correlated with increased cell death after introducing DNA DSBs after X-Ray irradiation. I will also investigate whether *MERIT40* silencing is causing an increased accumulation of spontaneous DNA DSBs by using immunoconjugates that target γH2AX to monitor DNA damage.
3. Evaluate the chemoresistance to cisplatin and carboplatin of EOC cell lines with silenced *MERIT40* compared to the parental EOC cell lines.
4. Investigate whether *MERIT40* knockdown in EOC cell lines would affect:
   A) cell cycle progression using FACS analysis
   B) anchorage dependent proliferation using MTT assays
   C) anchorage independent growth using soft agar assays
   D) the ability of cells to migrate using transwell migration assays
5. Evaluate the expression of genes involved in cell cycle progression, proliferation, apoptosis, proliferation, chromosomal segregation and DNA repair pathways after *MERIT40* knockdown in EOC cell lines, by performing a Fluidigm gene expression experiment.

6.2 Evaluation of *MERIT40* knockdown effect in platinum response and DNA damage repair ability of EOC cell lines

Cisplatin binds to the DNA backbone and causes the formation of interstrand crosslink adducts. The physical distortions created on the DNA induced by those intrastrand platinum adducts are recognised by DNA damage recognition proteins which in turn are transducing DNA damage signals to downstream effectors and if DNA repair pathways are defected apoptosis is triggered. To evaluate the role of MERIT40 as part of the BRCA1 complex involved in HR triggered after DNA damage by platinum agents, my goal was to select EOC cell lines, knock down *MERIT40*, confirm the knockdown of the protein and monitor the cell death caused by cisplatin administration in the presence or absence of MERIT40.

6.2.1 Selection of EOC cell lines for *MERIT40* knockdown

In order to select at least two EOC cell lines to knockdown *MERIT40* and evaluate any differences in the chemoresistance I monitored a panel of highly expressing *MERIT40* EOC cell lines for their sensitivity to cisplatin (also known as CCDP), the first of two chemotherapeutic compounds I would use. Carboplatin was also used later.

6.2.1.1 Selection of highly expressing *MERIT40* EOC cell lines

Seven EOC cell lines that demonstrated higher *MERIT40* expression compared to NOSE in the Taqman Real time expression assay in the pilot study previously described were chosen to be evaluated for their chemoresistance to cisplatin. Additionally, I monitored two pairs of cell lines, A2780CP/A2780 and C13/OV2008 which are pairs of cisplatin resistant and sensitive strains of the same cell line. A2780 is the cisplatin sensitive strain and A2780CP is the cisplatin resistant strain of the same cell line, and OV2008 and C13 are the

cisplatin sensitive and resistant strains respectively of the same cell line. The histology of the selected EOC cell lines monitored is shown on Table 6.1.

Because no expression data were available for the cisplatin resistant and sensitive pairs and in order to confirm the elevated expression of *MERIT40* compared to NOSE, the 11 cell lines were cultured again and RNA was extracted, reverse transcribed and used in a validating *MERIT40* Taqman Real time expression assay. The expression of *MERIT40* in the selected EOC cell lines was compared to a small number of NOSE cell lines (Figure 6.1).

Each of the 11 EOC cell lines exhibited at least a two-fold higher *MERIT40* expression than NOSE cell lines, thus confirming that they were all suitable to take forward in *MERIT40* knockdown experiments. *MERIT40* expression was compared the EOC and NOSE cell line by using the Wilcoxon Rank sum test. *MERIT40*'s significant over-expression in EOC cell lines was confirmed with no overlap observed between NOSE and EOC cell lines (Figure 6.2)

EOC cell line	Histology
TOV112D	Endometrioid
EFO27	Mucinous papillary adenocarcinoma
HOC7	Well differentiated serous, ascites
MPSC1	Low-grade serous carcinoma
OAW42	Serous cystadenocarcinoma in relapse
OVCAR3	High grade serous adenocarcinoma, from ascites
SCOV3IP	Moderately well differentiated adenocarcinoma from ascites, from xenograft
A2780	Undifferentiated adenocarcinoma
A2780CP	Cisplatin resistant undifferentiated carcinoma
OV2008	Endometrioid with squamous differentiation, from xenograft
C13	Cisplatin resistant endometrioid with squamous differentiation

Table 6.1: Histological information for the EOC cell lines that were monitored for chemoresistance to cisplatin.

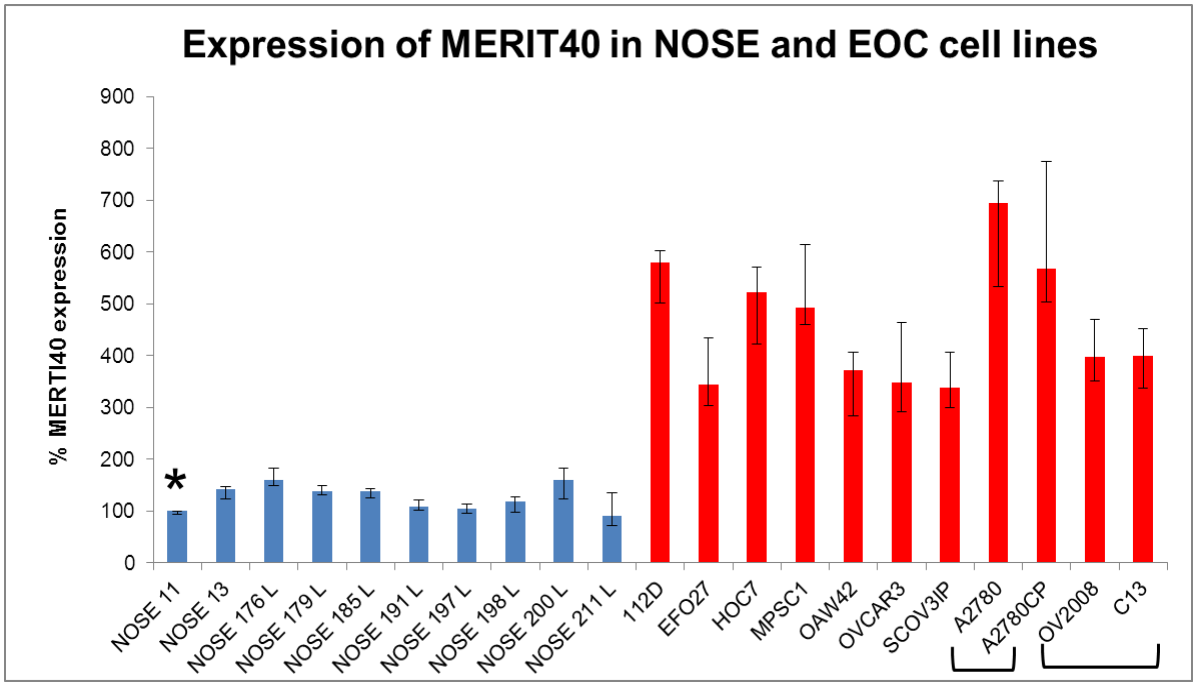


Figure 6.1**:** Differential *MERIT40* expression between the 11 selected EOC cell lines and 10 randomly selected NOSE cell lines normalised to β-actin. (*): NOSE11 *MERIT40* expression was used as the calibrator sample and the expression of the rest of the cell lines was calculated relative to that sample using the ΔΔCt method. A2780-A2780CP and OV2008-C13 are the cisplatin sensitive-resistant pairs.

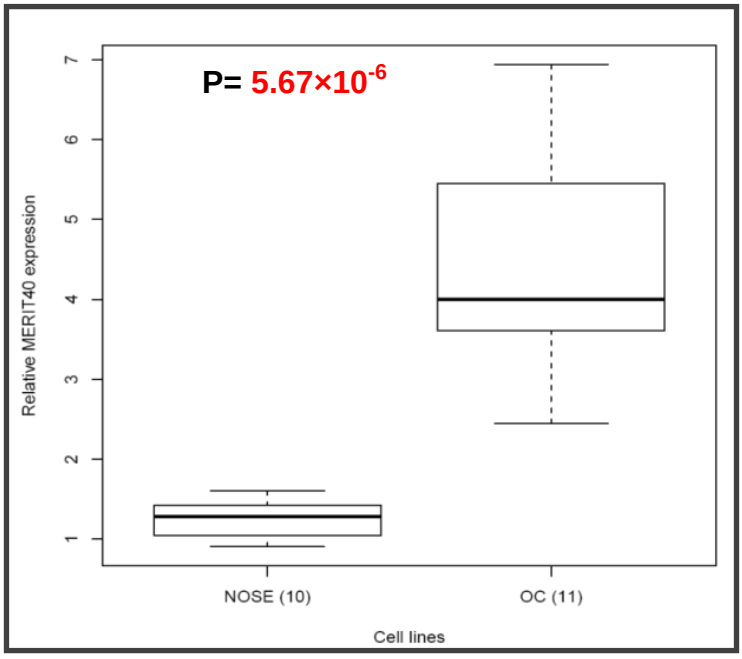


Figure 6.2: Wilcoxon rank sum test for *MERIT40* expression in NOSE versus the selected 11 EOC cell lines

6.2.1.2 Cisplatin dose-response curves for the EOC cell lines with high *MERIT40* expression

The cisplatin sensitivity for each of the selected cell lines was assessed by monitoring their viability in response to dosing with increasing concentrations of cisplatin after 72 hours. The viability of cell lines following cisplatin dosing was measured by performing an MTT assay as described in the methods section. The effectiveness of cisplatin to cause death to each cell line was deduced through the production of dose response curves. From these curves the IC50 (the half maximal inhibitory concentration) of cisplatin which translates to the amount of cisplatin needed (dose) for 50% of the cells to die (response) was calculated for each of the selected EOC cell lines.

Cisplatin dosing and MTT assays were performed for the selected EOC cell lines as described in the methods section. Cell line 112D solubilised after addition of MTT so it was excluded from the analysis. Each condition was repeated in quadruplicate and each assay was repeated twice. In order to calculate the cisplatin IC50 for each cell line, GraphPad software was used by which non-linear regression with variable slope analysis was performed. The dose response curves generated for each cell line and the IC50s generated are shown in Figure 6.3 and Table 6.2 respectively.

The cisplatin IC50 for the sensitive-resistant cell lines A2870 and A2780CP was found to be 0.19 μM and 8.43 μM respectively, showing a 44 fold increased resistance to cisplatin for A2780CP. For the other pair, OV2008 and C13, the IC50s were found to be 0.17 μM and 36.2 μM respectively, showing a 213 fold increased cisplatin resistance for C13. A2780CP is the resistant counterpart of A2780 cell line, but it demonstrated almost ten times less resistance to cisplatin than C13, the resistant counterpart of OV2008. The cisplatin IC50s for the remaining cell lines ranged from 2.62 μM for OAW42 to 19.28.μM for OVCAR3. Cell lines OAW42, SCOV3IP and EFO27 appeared to be more sensitive to cisplatin whereas MPSC1, HOC7 and OVCAR3 appeared fairly resistant.

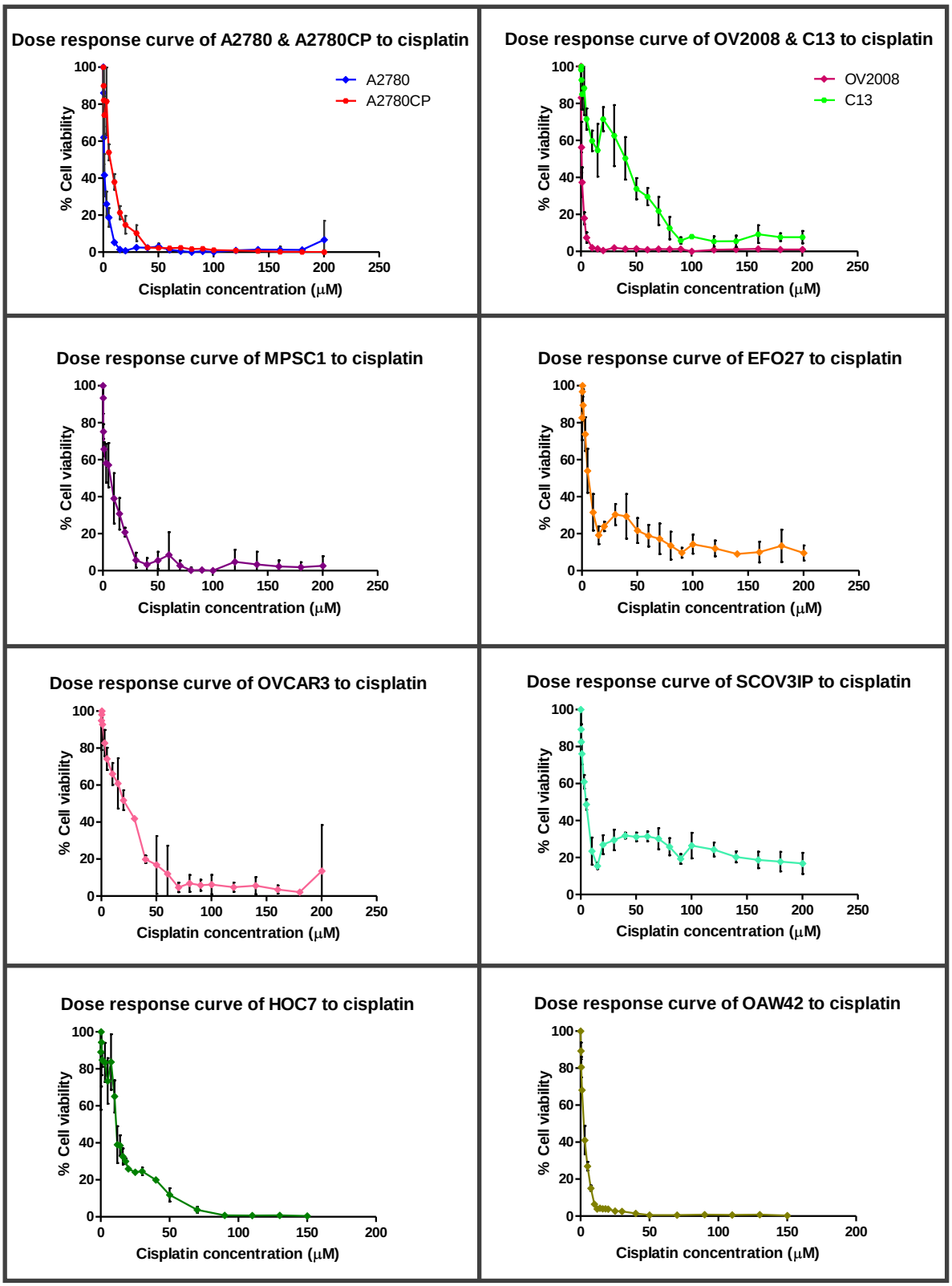


Figure 6.3**:** Dose response curves for EOC cell lines after dosing with cisplatin. Viability of cells was measured by performing an MTT assay. The difference in IC50s in the two sensitive-resistant pairs is obvious from the dose response curves at the top of the figure. The steeper the curve is the more efficiently cisplatin kills the cells, thus the more sensitive the cell line.

EOC cell line	Cisplatin IC50 (μM)
TOV112D	Detached with MTT addition
OAW42	2.62
SCOV3IP	3.78
EFO27	4.51
MPSC1	8.35
HOC7	12.7
OVCAR3	19.28
A2780	0.19
A2780CP	8.43
OV2008	0.17
C13	36.2

Table 6.2**:** Cisplatin IC50s for the selected EOC cell lines. Highlighted in purple are the EOC cell lines with high expression of *MERIT40* listed in order of increasing resistance to cisplatin. Highlighted in yellow and green the two platinum sensitive-resistant pairs.

The purpose for assaying the resistance of these highly expressing *MERIT40* cell lines to cisplatin was to select the appropriate cell lines for stably knocking down *MERIT40* and reassessing their resistance to the drug. I selected cell lines that would not be over-sensitive to cisplatin because I wanted to have a large enough window to observe decreased resistance to cisplatin after knocking down *MERIT40*. That ruled out the two cell lines that showed the lowest sensitivity to cisplatin, OAW42 and SCOV3IP. Additionally, due to the heterogeneity in ovarian cancer, I decided to carry the knockdown of *MERIT40* in cell lines representing different subtypes. According to these selection criteria, the cell lines that were selected for *MERIT40* knockdown were EFO27 (Mucinous, Cisplatin IC50= 4.52), MPSC1 (Low grade serous, Cisplatin IC50= 8.35), OVCAR3 (High grade serous, Cisplatin IC50= 19.28) and A2780CP (Undifferentiated, Cisplatin IC50= 8.43). A summary of the features of the cell lines selected to knockdown *MERIT40* is shown in Table 6.3.

EOC cell line	Histology	Cisplatin IC50 (μM)	95% CI
EFO27	**Mucinous papillary adenocarcinoma**	4.51	3.44- 5.9
MPSC1	**Low grade serous carcinoma**	8.35	6.47- 10.77
OVCAR3	**High grade serous adenocarcinoma**	19.28	15.18- 24.47
A2780CP	**Undifferentiated adenocarcinoma**	8.43	7.12- 9.98

Table 6.3: EOC cell lines selected for *MERIT40* knockdown. EOC cell lines represent four different subtypes of EOC and they demonstrated a cisplatin IC50 large enough to provide a window to observe decrease in cisplatin resistance after *MERIT40* was knocked down.

6.2.2 Generating EOC cell lines with stable MERIT40 knockdown

According to the real time expression data for *MERIT40* expression and the cisplatin IC50s generated by dose response curves, four cell lines were selected to perform stable knockdown of *MERIT40*. The four EOC cell lines selected (EFO27, MPSC1, OVCAR3 and A2780CP) were infected with lentiviruses containing an empty GIPZ plasmid or GIPZ with non-silencing shRNA as negative controls, shRNA against GAPDH as positive control to assess the effectiveness of the viruses, and three different shRNAs targeting *MERIT40* (M1, M2, M3). The plasmids all contained a GFP (Green Fluorescent protein) gene and a puromycin gene for selection. After large scale preparation of the plasmids, they were sequenced to confirm their sequence targeting the intended *MERIT40* loci. Plasmids M1 and M3 demonstrated the right hairpin sequences but plasmid M2 had a base substitution. However, infection with a virus containing the M2 plasmid was performed anyway as this base substitution might not influence the effectiveness of the shRNA to silence *MERIT40*. The infection was performed at MOI=1 (multiplicity of infection meaning lentiviral units per cell) after titration of each virus as described in the methods section to establish the TU for each virus. To ensure that 100% infected cells would be used in my subsequent phenotypic assays selection with puromycin was performed until the relative control-uninfected cell line was killed off. The infected cell lines were constantly kept under puromycin selection in subsequent culturing. The stable *MERIT40* knockdown cell lines were then used to generate RNA and protein to assess the knockdown levels of *MERIT40*

at the mRNA and protein level respectively. The process of generating the stable *MERIT40* silenced EOC cell lines is presented in Figure 6.4.

The next step was to check the *GAPDH* expression of the cell lines that were infected with lentivirus containing the shRNA against *GAPDH*. The expression of *GAPDH* in infected cell lines was calibrated to the expression of the Empty GIPZ infected cell line. The best *MERIT40* knockdown was achieved in EFO27 cell line with 74% *GAPDH* knockdown and in MPSCI with 47% knockdown. Cell lines A2780CP and OVCAR3 demonstrated only 20-24% knockdown indicating that they were more difficult cell lines to stably infect. The expression levels of *GAPDH* in the different cell lines are shown in Figure 6.5.

The level of *MERIT40* knockdown on the mRNA level was checked in all four cell lines (Figure 6.6). Construct M1 appeared to be the one that caused the most efficient knockdown of *MERIT40* after analysing mRNA expression levels of the gene by Taqman. The M3 shRNA caused partial *MERIT40* knockdown but not more than two-fold in any of the four cell lines apart from EFO27_M3. No *MERIT40* knockdown was achieved with the M2 shRNA so the cell lines infected with this virus were not included in subsequent experiments.

To ensure that the observed decrease in *MERIT40* mRNA correlated with depleted MERIT40 protein, I then generated protein lysates for all the infected and control cell lines. The western blots were generated using those lysates in USC (University of South California) by Dr Kate Lawrenson using a custom made antibody. The protein levels confirmed the gene silencing observed by the Real time gene expression assay on the mRNA level, similarly indicating that the most efficient knockdown was caused by *MERIT40* shRNA M1 (Figure 6.6, B).

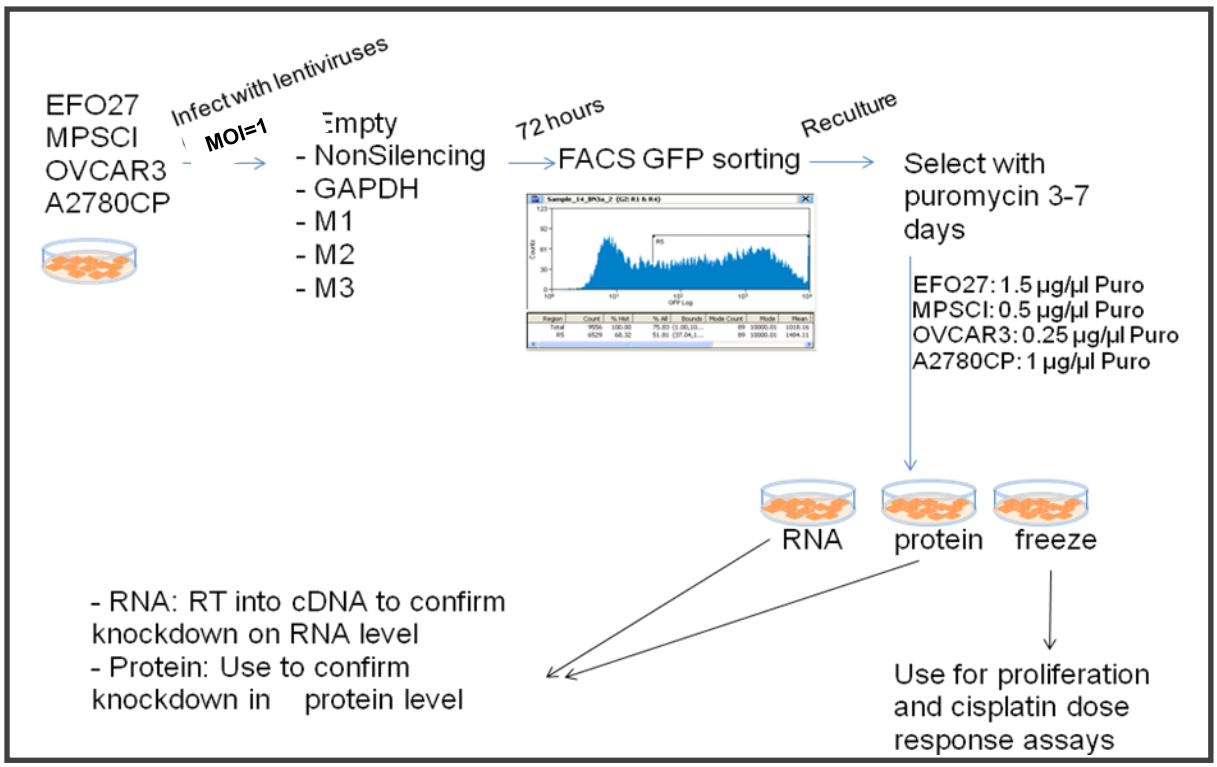


Figure 6.4**:** Generation of four EOC cell lines with stably silenced *MERIT40.* The generation of the stable cell lines was achieved using lentiviruses containing GIPZ-shRNA plasmids.

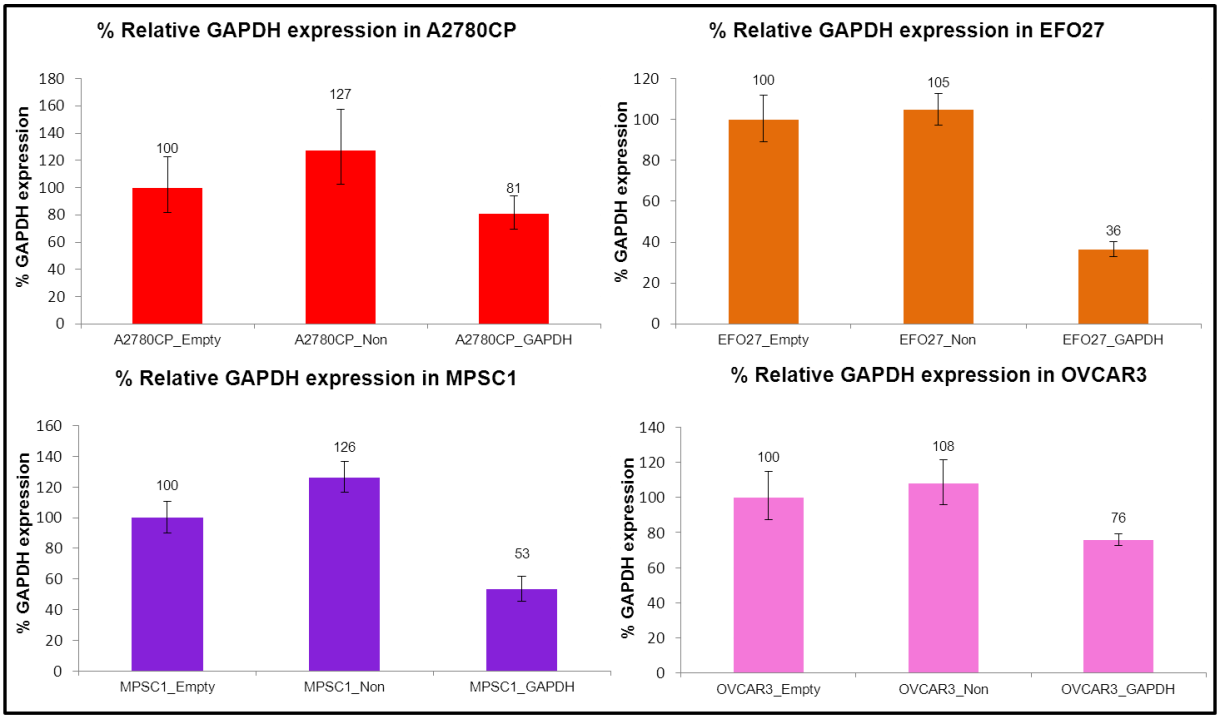


Figure 6.5**:** *GAPDH* expression of infected EOC cell lines. The histograms represent *GAPDH* expression normalised to β-actin and calibrated against the expression of the Empty GIPZ infected cell line. EFO27 and MPSC1 showed the most efficient knockdown but knockdown of *GAPDH* in A2780CP and OVACR3 was not as successful.

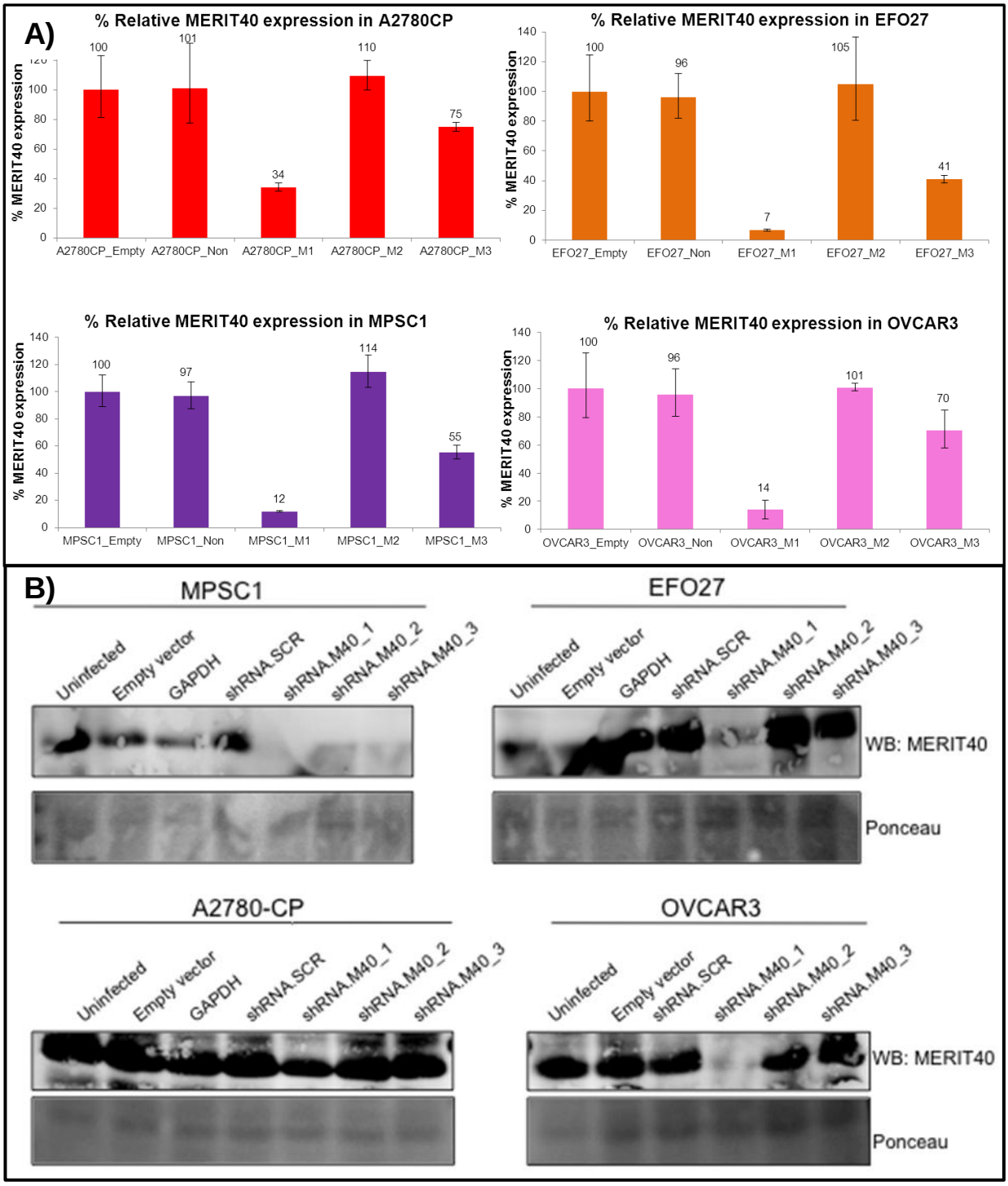


Figure 6.6**:** Confirmation of *MERIT40* knockdown in stably infected EOC cell lines. A) *MERIT40* knockdown of the stably infected EOC cell lines confirmed on the mRNA level. The histograms show *MERIT40* expression normalised to β-actin and calibrated against the expression of the Empty GIPZ infected cell line. All four EOC cell lines showed the most efficient knockdown with *MERIT40* shRNA M1. Partial knockdown of *MERIT40* was achieved with *MERIT40* shRNA M3. B) Western blots showing MERIT40 protein expression in control and *MERIT40* knockdown EOC cell lines. Ponceau staining was used as loading control. All four EOC cell lines showed the most efficient knockdown with *MERIT40* shRNA M1.

6.2.3 Response of EOC cell lines to chemotherapeutic agents following MERIT40 knockdown

The hypothesis behind this was that with MERIT40 involved in the BRCA1-Rap80 complex, cells with absent MERIT40 would have impaired HR machinery and would be unable to repair DNA DSBs induced by platinum agents and ionising radiation, thus increased cell death would be observed. The control and the *MERIT40* shRNA M1 and M3 stably infected cell lines were then used to perform dose response curves and generate IC50s in order to assess if there was a potential role of MERIT40 in conferring chemoresistance. I performed the dose response assays in duplicate for all cell lines. The chemotherapeutic agents that I used were cisplatin (CCDP) and carboplatin and the cells were subjected to dosing of cisplatin for 72 hours (Figure 6.8) and carboplatin for 96 hours (Figure 6.9). Carboplatin dosing was also performed for 72 hours but the doses required were extremely high and even in the highest dosing the 100% cell death could not be achieved in any of the cell lines at the 72 hour timeframe.

Only EFO27 with M1 (EFO27_M1) *MERIT40* knockdown showed a minor decrease (less than two-fold) in the cisplatin IC50 compared to the EFO27_Empty negative control after dosing for 72 hours. The rest of the cell lines with *MERIT40* knockdown did not show any increased sensitivity to cisplatin. None of the cell lines showed increased resistance to carboplatin when *MERIT40* was knocked down. The summarised data of the dose response assays to cisplatin and carboplatin for all the cell lines are shown in Table 6.4 and Figure 6.7.

Cell lines		Cisplatin IC50 (μM) [(95% CI]_72 hours	Carboplatin IC50(μM) [(95% CI]_96 hours
EFO27	**EFO27**	7.96 [7.39-8.58]	17.26 [14.40-20.70]
	EFO27_Empty	9.92 [9.45-10.42]	32.25 [25.65-40.54]
	EFO27_Non	8.69 [8.06-9.37]	29.44 [20.95-41.36]
	EFO27_M1	**5.34 [4.55- 6.28]**	39.8 [35.57- 44.54]
	EFO27_M3	7.63 [6.41-9.09]	32.31 [25.42-41.06]
MPSC1	**MPSC1**	15.47 [14.77-16.19]	48.94 [42.76-56.01]
	MPSC1_Empty	15.58 [14.91-16.28]	58.43 [43.23-78.97]
	MPSC1_Non	15.06 [14.49-15.64]	42.25 [35.57-50.20]
	MPSC1_M1	13.94 [13.37-14.53]	53.48 [34.76-82.27]
	MPSC1_M3	14.28 [13.62-14.97]	44.25 [40.47-48.40]
A2780CP	**A2780CP**	6.81 [5.81-7.98]	72.72 [54.67-96.73]
	A2780CP_Empty	10.13 [8.822 -11.63]	62.22 [55.83-69.35]
	A2780CP_Non	9.81 [8.88-10.83]	57.92 [53.03-63.25]
	A2780CP_M1	10.76 [9.76-11.87]	52.2 [43.42-62.74]
	A2780CP_M3	7.85 [6.85-9.01]	79.62 [60.04-105.6]
A2780	**A2780**	0.41 [0.052-3.19]	12.47 [11.11-14.00]
OVCAR3	**OVCAR3**	20.25 [17.8-23.04]	22.94 [19.28-27.28]
	OVCAR3_Empty	23.53 [18.51-29.92]	59.36 [11.29-312]
	OVCAR3_Non	29.01 [24.46-34.3]	62.25 [11.46-338.3]
	OVCAR3_M1	31.91 [28.77-35.41]	18.84 [15.24-23.30]
	OVCAR3_M3	43.52 [38.52-49.43]	38.91 [24.53-61.71]

Table 6.4**:** Cisplatin and carboplatin IC50s and 95% confidence intervals (CI) for the selected EOC control and *MERIT40* knockdown cell lines.

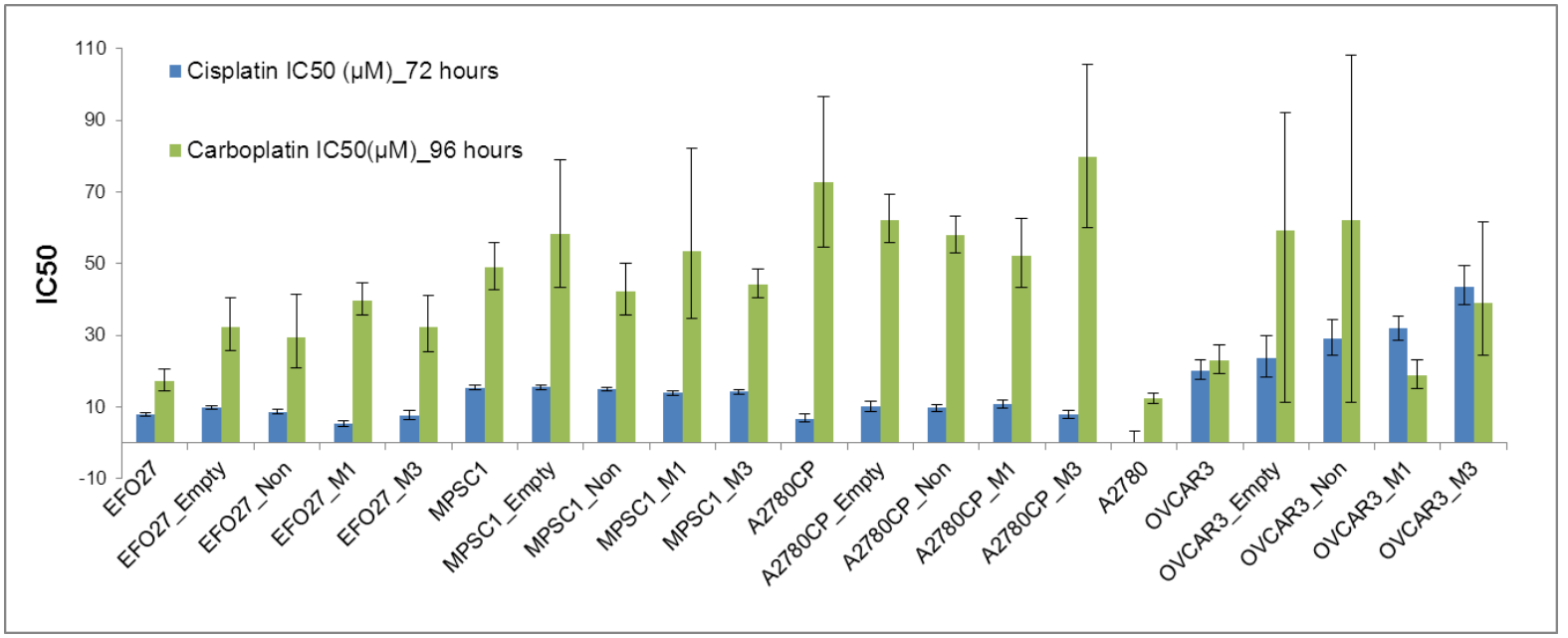


Figure 6.7**:** Histogram showing the IC50s of all control and *MERIT40* knockdown cell lines for the two chemotherapeutic drugs tested, cisplatin and carboplatin. There were no cell lines that showed a twofold or higher decrease in the IC50 when *MERIT40* is knocked down. The error bars are representing the 95% confidence intervals for the calculated IC50s as calculated in PRISM using non-linear regression.

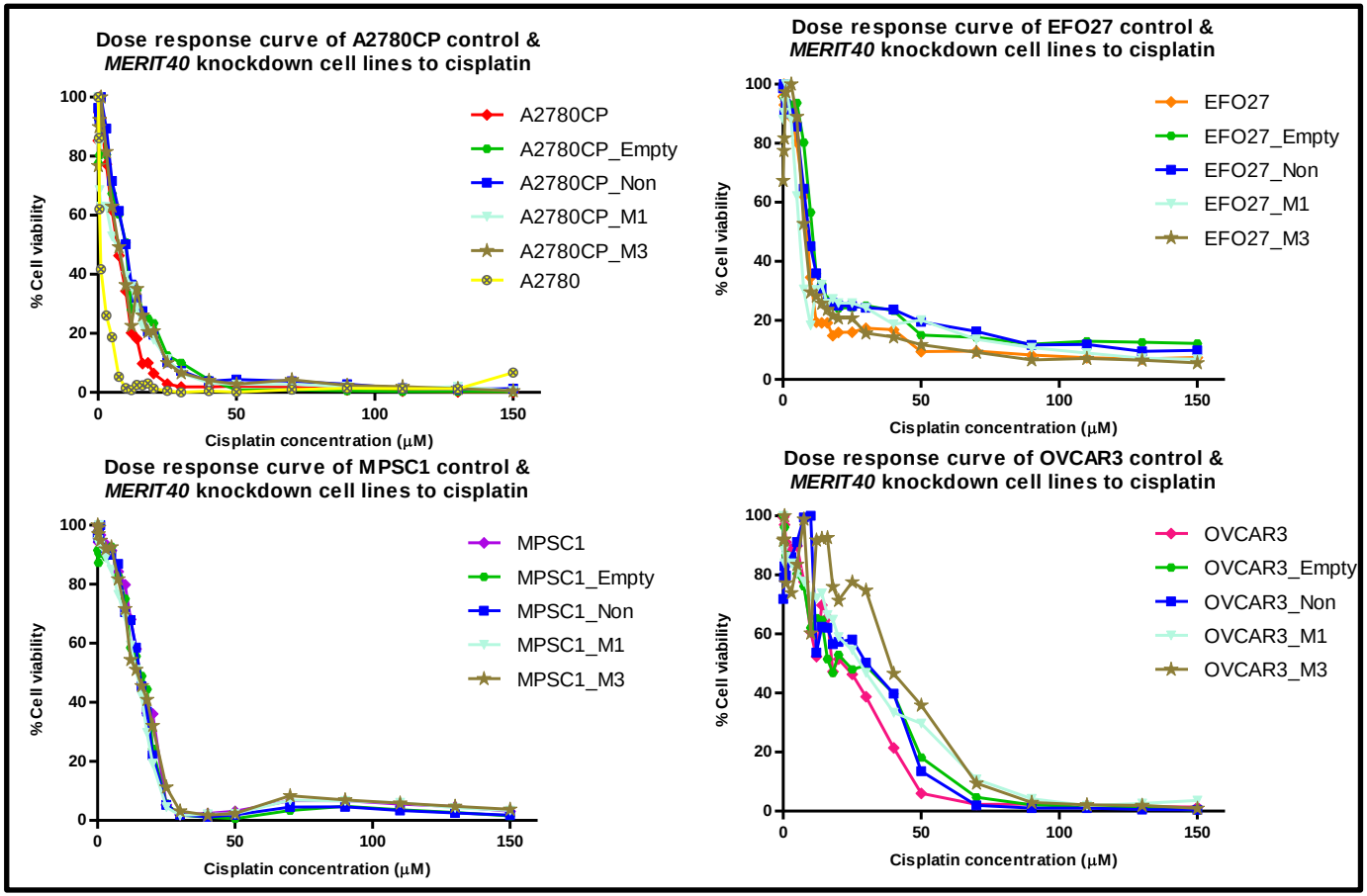


Figure 6.8**:** Dose response curves of EOC control and *MERIT40* knockdown cell lines after cisplatin dosing. Viability of the cells was measured by performing an MTT assay 72 hours later. For the A2780CP derived stable cell lines the cisplatin sensitive counterpart A2780 was used as a control and showed the expected increased sensitivity to cisplatin compared to A2780CP.

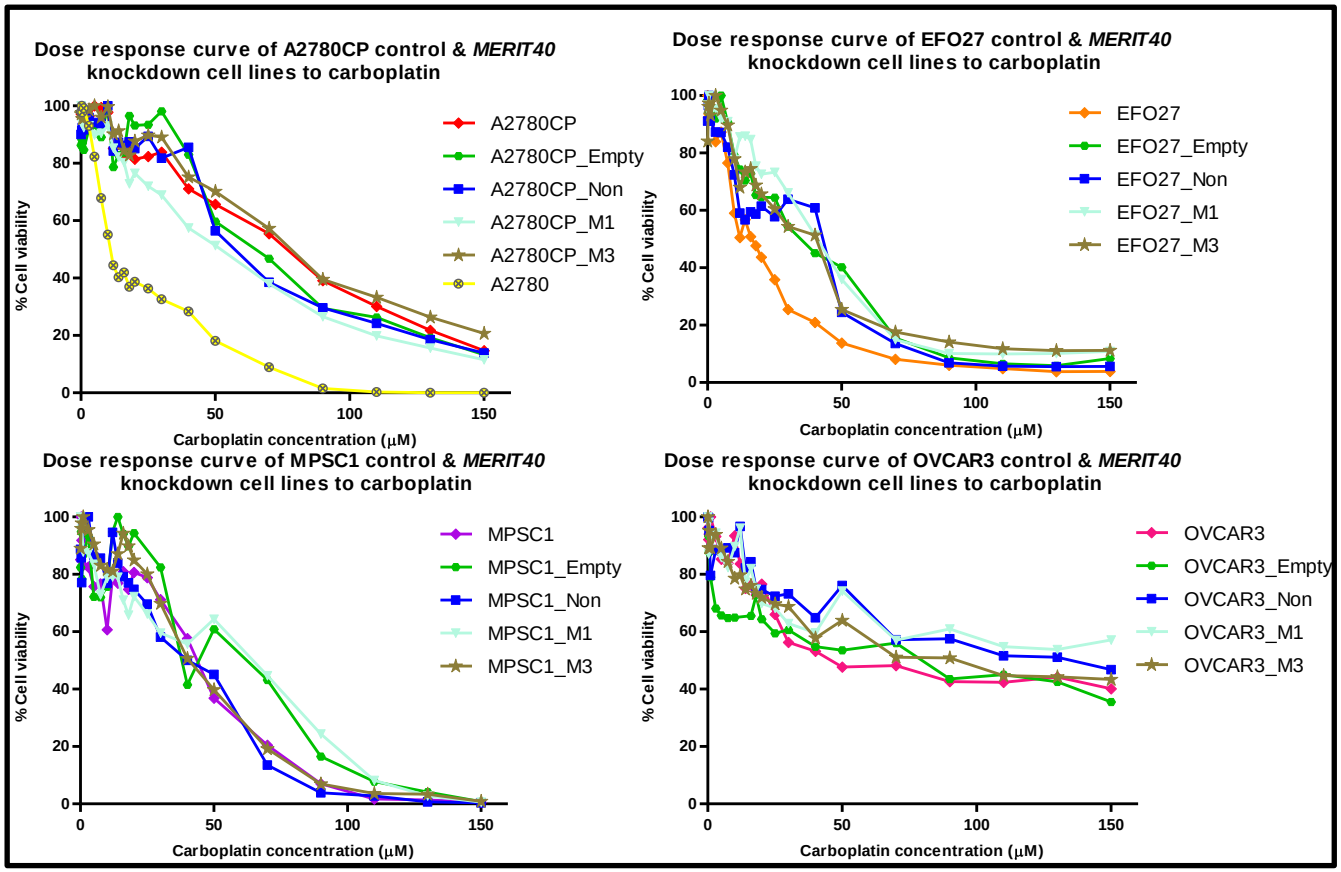


Figure 6.9**:** Dose response curves of EOC control and *MERIT40* knockdown cell lines after carboplatin dosing. Viability of the cells was measured by performing an MTT assay 96 hours later. A2780 showed the expected increased sensitivity to carboplatin compared to A2780CP.

6.2.4 Evaluating the effect of silencing *MERIT40* in EOC cell lines on their efficiency to repair DNA DSBs caused by endogenous or exogenous stress.

It has been recently shown that MERIT40 is recruited to sites of DNA DSBs resulting from endogenous and exogenous genotoxic stress and is conferring resistance to HeLa cells subjected to ionising radiation by facilitating the association of the deubiquinating agent BRCC36 with BRCA1-Rap80 and the localisation of the complex at the DNA DSBs (Shao et al, 2009). I have evaluated whether the controls and *MERIT40* knockdown EFO27 EOC cell lines show any difference in sensitivity to ionising radiation. I have applied increasing concentrations of irradiation (IR measured in Gray) and measured the cell viability after 72 hours in culture by performing an MTT assay. The % of cell viability for EFO27 cells when *MERIT40* was knocked down compared to the controls decreased after exposure to increasing doses of ionising radiation as shown in Figure 6.10. The IC50s for each cell line in response to IR were calculated using Prism performing a logIR versus response dose response inhibition analysis. I observed an at least 2 fold decrease in the IC50s for EFO27_M1 and EFO27_M3 compared to the controls indicating that depleting MERIT40 sensitised the ovarian cancer cells to DNA damage caused by irradiation (Table 6.5). However, it should be noted that the 95% confidence intervals calculated for the IC50s where quite large so the observed result should be treated with caution.

Additionally, to find out just how essential the role of MERIT40 is in DNA repair in EOC cell lines, I investigated whether after knocking down *MERIT40* the cells would be more susceptible to spontaneous DNA damage under normal culturing conditions. As a marker I used phosphorylated γH2AX which localises in the nucleus in response to DNA damage and its presence would be indicative of accumulated unrepaired DNA damage. After culturing of the cell line for 48 hours I found an increase on the localisation of phosphorylated γH2AX in the nuclei of the EFO27_M1 cell line. Cells were sparsely seeded on coverslips and cultured for 48 hours, then fixed and immunofluorescence was performed using an antibody against phosphorylated γH2AX. Following the staining, 80-100 cells were captured and the cells with no, low or high γH2AX expression were

counted. The % of cells with high γH2AX expression was increased by 20% (2 fold increase) when MERIT40 was absent, a difference mainly attributed to a significant reduction in the number of cells with no γH2AX and a smaller reduction in the number of cells with low γH2AX (Figure 6.11). This observation indicates that when MERIT40 is absent the accumulation of unrepaired spontaneous DNA damage is increased.

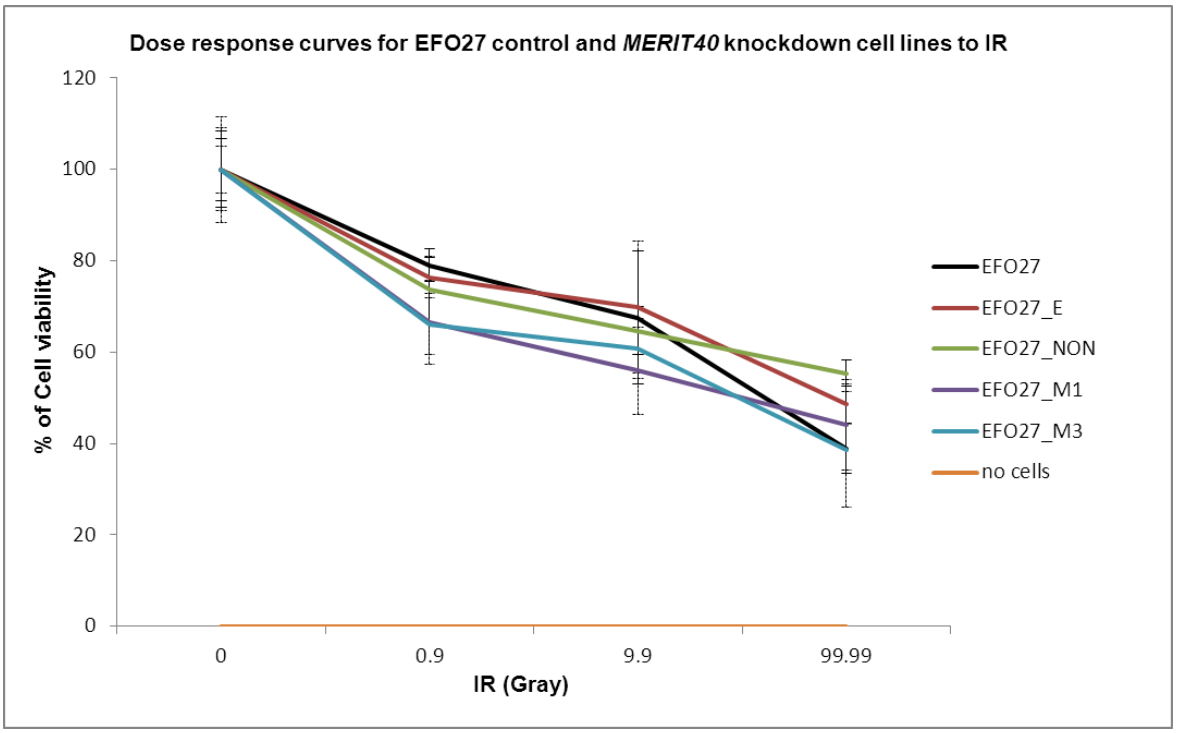


Figure 6.10: Dose response curves for EFO27 *MERIT40* knockdown and control cell lines after subjecting them to increasing doses of X-Ray. The % of cell viability was measured by measuring the absorbance of cells after the addition of MTT and DMSO after 72 hours of incubation.

Cell line	IR IC50 (Gray) [(95% CI]_72 hours
EFO27	40.29 [27.52-58.99]
EFO27_Empty	69.02 [38.57-123.5]
EFO27_Non	70.32 [46.51-106.3]
EFO27_M1	**26.86 [16.27-44.36]**
EFO27_M3	**15.89 [10.24-24.66]**

Table 6.5: Calculated IC50s for the EFO27 control and *MERIT40* knockdown cell lines in response to increasing doses of irradiation. 95% CI: 95% confidence intervals.

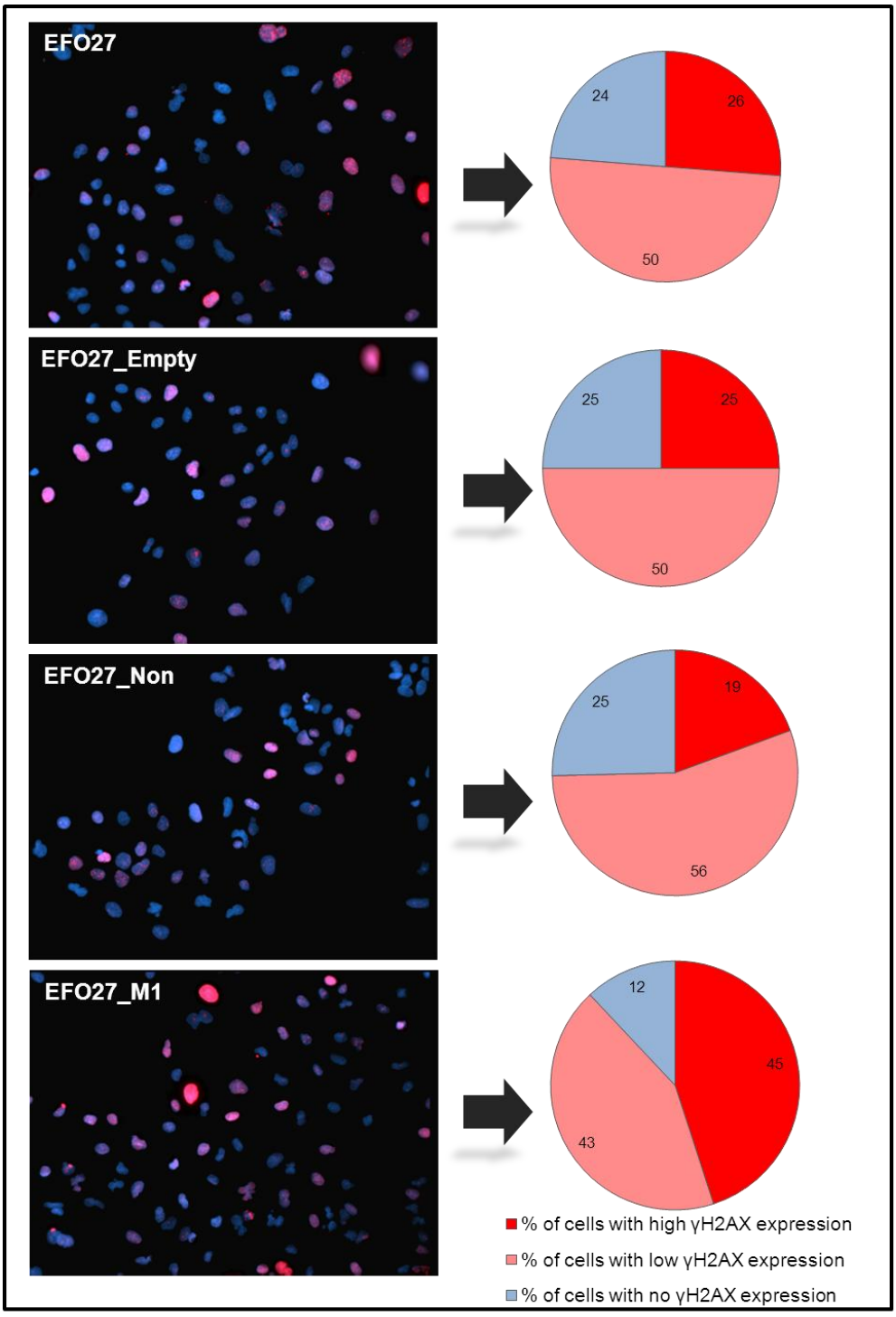


Figure 6.11: Immunofluorescent staining and quantification of phosphorylated γH2AX in EFO27 control and *MERIT40* knockdown cells. Staining for γH2AX was performed in order to evaluate the cell lines' efficiency for DNA repair in response to spontaneous endogenous damage. The blue colour is DAPI staining for visualisation of the cell nuclei. The pink staining represents the γH2AX expression in the nuclei. The corresponding pie charts are representing in blue the % of cells with no γH2AX expression, in pink and red the % of cells with low and high γH2AX expression respectively.

6.3 Phenotypic analysis following *MERIT40* knockdown in EOC cell lines

MERIT40 was found to be over-expressed in EOC cell lines indicating a gain of function role in ovarian cancer development. This may involve function of MERIT40 in other pathways than the DDR pathway. For the development of neoplastic transformation leading to malignant tumorigenesis cells are subjected to several phenotypic changes ranging from unregulated cell cycle progression, proliferation in anchorage independent and anchorage dependent manner, increased ability of cells to migrate or invade through extracellular matrix. The elevated expression of *MERIT40* in EOC compared to NOSE cells was a significant observation and I decided to investigate further whether some of the phenotypic changes indicative of neoplastic transformation were linked to the elevated expression of this gene. By knocking down *MERIT40* in EOC cell lines and assaying the parental and *MERIT40* silenced cell lines for reversal of neoplastic phenotype I aimed to shed some light in the role of the gene in ovarian cancer development.

6.3.1 Investigating the potential role of MERIT40 in cell cycle progression and proliferation

The process of cell division by which a single cell forms two daughter cells is divided into G1, S and G2/M phases that collectively constitute the cell cycle. The phases of the cell cycle are distinguished according to the amount of DNA a single cell has. The cells in G1 phase have a 2N DNA content, then during the S phase the content of DNA steadily increases until the G2 phase when the cell has 4N DNA content where completely duplicated DNA is reached. Each phase is characterised by certain activities that are required for proliferation to proceed to the next phase. During cancer development, activation of oncogenes has been shown to regulate the progression of the cell cycle by promoting cell proliferation.

The gain-of-function role of *MERIT40* attributed by over-expression of the gene in EOC cell lines compared to NOSE & FTE cell lines might be translated in an involvement of MERIT40 in regulating cell cycle progression and

proliferation. Parental, control and *MERIT40* silenced cell lines were cultured in duplicates, harvested and stained with a DNA intercalating agent Propidium Iodide (PI) that indicates the DNA content of the cells. Cell cycle profiles were generated by FACS analysis to investigate any effect *MERIT40* knockdown may have on the original EOC cell line's cell cycle. The analysis was performed using the Summit v4.3 software. An approximate 15-20% decrease of the cells in G2/M phase was observed for cell line EFO27_M1 and EFO27_M3 in the duplicate experiments (Figure 6.12). No cell cycle arrests were observed for any of the other EOC cell lines (Figure 6.12). An interesting observation is the 50% decrease in cells with 8N content found for cell lines EFO27_M1 and EFO27_M3. The cell cycle profiles generated by FACS clearly show the cell cycle arrests observed (Figure 6.13).

Cells were arrested in G1 phase when MERIT40 was not expressed for EFO27. A reduction in the cells with 8N DNA content may suggest that the arrest observed was not due to reduced proliferation but due to reduction in ploidy. No differences in the cell cycle profiles were observed for the other EOC cell lines with silenced *MERIT40* (Appendix, Figure 1). The histological diversity of ovarian cancer is reflected in the EOC cell lines selected to knock down *MERIT40* and it should come as no surprise that different molecular changes might have contributed to neoplastic development across the different subtypes.

A cell that appears to have a 4N content could be a cell in G2 phase but could also be a tetraploid cell on G1 phase or even an anaploid cell that has abnormal chromosomes. Often cancer cell lines are not diploid, as normal cell lines are, but tetraploid, anaploid or octaploid. In fact EFO27 has been shown to be hypotetraploid (Lambros et al, 2005). I prepared DNA metaphase spreads of EFO27 and its *MERIT40* knockdown derivatives to examine if there was any reduction in ploidy by counting the chromosomes. I found that more cells in EFO27 appeared to be polyploid (with 4N or 8N DNA content) and found no polyploid cells in the EFO27_M1 cell line metaphase preparations (Figure 6.14). The number of cells available to be analysed in the metaphase preparations was too small to assess statistical significance of the observation. Thus, the absence of polyploid cells could be due to chance.

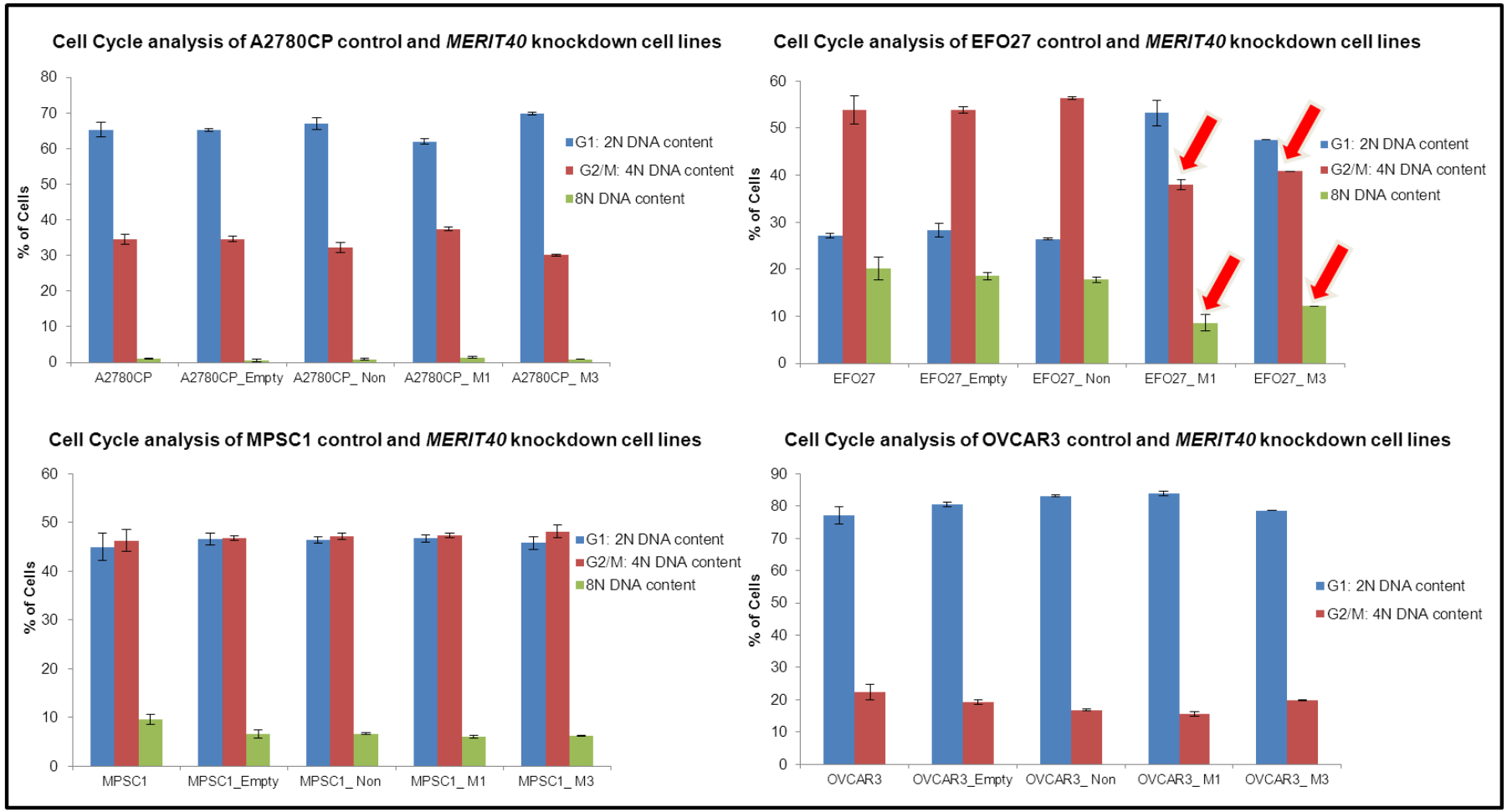


Figure 6.12**:** Histograms representing the amount of *MERIT40* knockdown and control EOC cells in the different cell cycle phases. Cell cycle analysis was performed using PI and the amount of cells in each phase was quantified using the Summit 4.3 software. The red arrows are showing the decrease of cell accumulation observed for the G2/M phase and for cells with 8n DNA content.

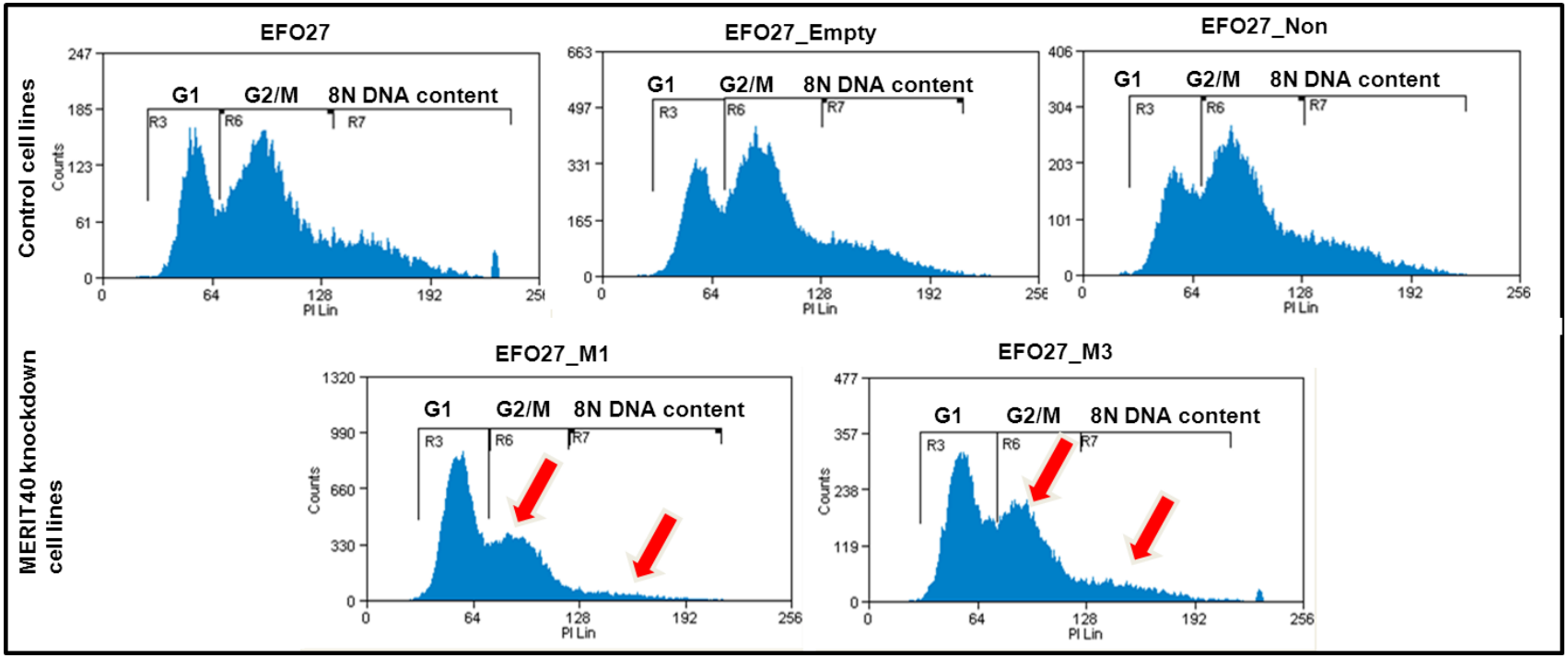
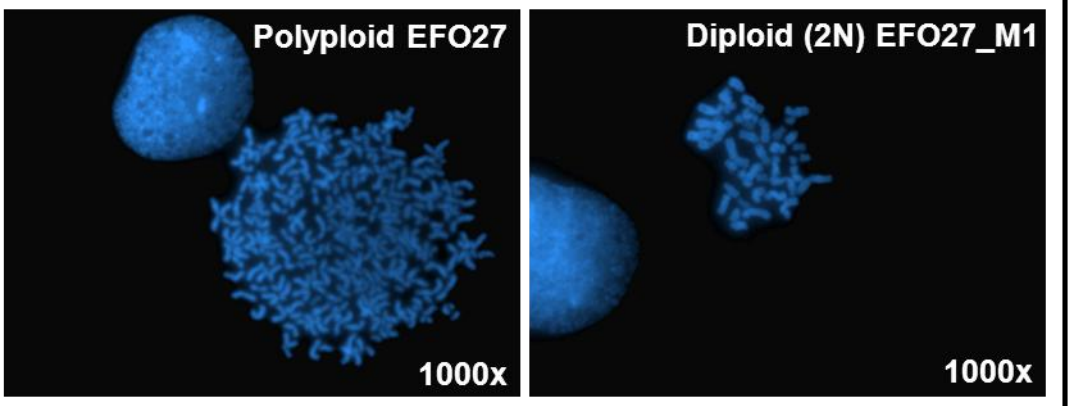


Figure 6.13**:** Cell cycle profiles of EFO27 control and *MERIT40* knockdown cell lines. The y axis corresponds to the counts, meaning the numbers of cells counted, and the x axis corresponds to the DNA content which is measured by the amount of PI incorporation in the DNA of the live single cells counted. The first peak in each profile, annotated as R3, indicates cells that are in G1 phase of the cell cycle, with a 2N DNA content. The second peak in each profile, annotated as R6, indicates cells that are in G2/M phase of the cell cycle with a 4N DNA content. The areas annotated as R7, correspond to cells with 8N or higher DNA content.

A)



B)

	EFO27	EFO27_Empty	EFO27_Non	EFO27_M1	EFO27_M3
Cells analysed	8	5	6	4	5
Diploid cells	6	4	4	4	5
Polyploid cells	2	1	2	0	0

Figure 6.14**:** Metaphase spreads and chromosome count of EFO27 control and *MERIT40* silenced cell lines. A) Representative pictures of DAPI stained metaphase spreads for one of the EFO27 control cells found to have abnormally high chromosome numbers (LEFT) and an EFO27_M1 cell found to have normal chromosome count (2N) (RIGHT). The pictures were obtained under the 100x oil objective lense of an immunofluorescent microscope. B) Table showing the numbers of diploid and polyploid cells found per cell line examined. No polyploid cells were observed in the EFO27 with knocked down *MERIT40.*

The proliferation of the EOC cell lines with active or silenced *MERIT40* was then assessed by performing an MTT assay as described in the methods section. The proliferation of EFO27_M1 and EFO27_M3 was found to be significantly reduced by ~4 fold and ~3 fold respectively compared to the parental EFO27 EOC cell line. The reduction was statistically significant in a two-tailed paired student t-test ($P= 3.4\times10^{-5}$ for EFO27_M1 and $P= P= 6.9\times10^{-5}$ for EFO27_M3). A significant 2 fold reduction was also observed in the proliferation of MPSC1_M1 ($P= 7.9\times10^{-4}$, Figure 6.15). An interesting observation was that in both cell lines the amount of proliferation reduction closely resembled the amount of *MERIT40* knockdown. In the case of EFO27

the reduction in proliferation was higher in EFO27_M1 that demonstrated a higher efficiency of *MERIT40* silencing compared to EFO27_M3 (Figure 6.6). *MERIT40* knockdown did not seem to cause any decrease in the proliferation of A2780CP and OVCAR3 EOC cell lines (Figure 6.15).

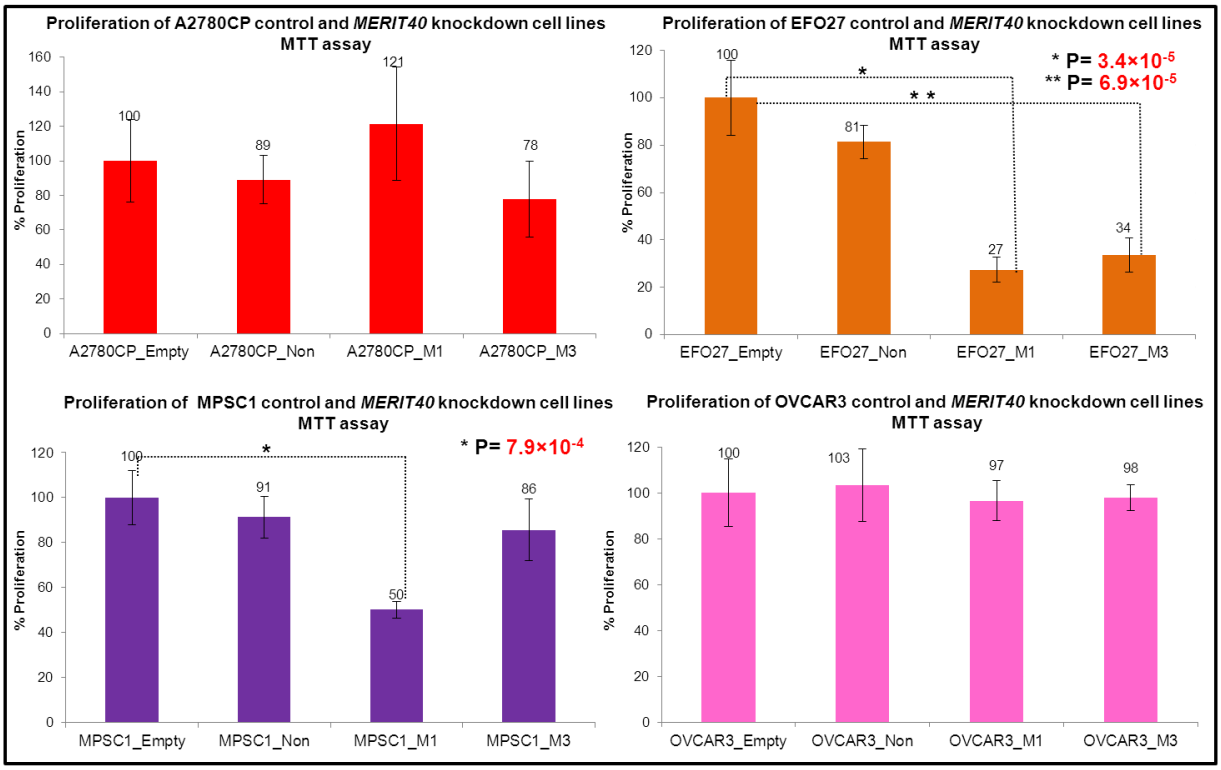


Figure 6.15**:** MTT proliferation assay of EOC control cell lines and *MERIT40* knockdown cell lines. The percentage of proliferation for each cell line was normalised against the proliferation of the cell line infected with a virus containing the empty GIPZ plasmid. A two tailed t-test was performed to calculate the significance of the observed decreases in proliferation and the P values are shown at the graphs.

Assessing the proliferation of cell lines using an MTT assay is a widely used and generally reliable method. However, it does not offer a precise "snap-shot" of the proliferative ability of a cell line such as would be provided by assessing the number of cells that are accumulated in the S phase of the cell cycle. I have previously performed a cell cycle analysis using PI staining, but the calculation of cells in S phase proved very difficult and inaccurate with just PI staining as the peak representative of S phase tends to fuse with G1 and G2/M phase peaks making the differentiation almost impossible and massively inaccurate.

Therefore, I decided to assess the proliferation ability of the cell lines with knocked down *MERIT40* by performing a Bromodeoxyuridine (BrDU) proliferation analysis that can provide a "snap-shot" window into the cell line proliferative ability by accurately quantifying the cells located on the S phase of the cell cycle. For this assay, BrDU was added to 80% confluent cultured cells for 1 hour and it incorporated into the newly synthesised DNA of replicating cells during the S phase of the cell cycle. The cells were fixed in ethanol and a FITC-labelled anti-BrDU antibody was used to stain all the proliferating cells. FACS analysis was used to quantify the amount of FITC stained cells, indicative of the cells in S phase.

Interestingly, the observations resulting from the BrDU assay contradicted the observations of reduced proliferation found by the MTT assay. The amount of cells with incorporated BrDU, indicative of them being in S phase, was not different between the parental and the *MERIT40* silenced cell lines for any of the four EOC cell lines tested. The FACS profiles and quantified cells in S phase for the parental and control EOC cell lines compared to their *MERIT40* knockdown counterparts are shown in Figures 6.16, 6.17, 6.18, 6.19. There might be several explanations for this outcome. The simplest one could be that the optimal BrDU incubation period may vary from cell line to cell line depending on how fast the cells go through the cell cycle to capture the cells entering S phase. The one hour incubation period used might not have provided a large enough window for changes to be captured. Further work would be required to investigate this. Alternatively, it could be that the BrDU proliferation assay actually provided the more accurate result. MTT is an assay that is used to assay proliferation as a first line assay but it does not really provide a snapshot in the proliferation, so the result observed could be due to increased apoptosis. Future work investigating the role of MERIT40 in apoptosis regulation would be interesting and could elucidate the conflicting observations.

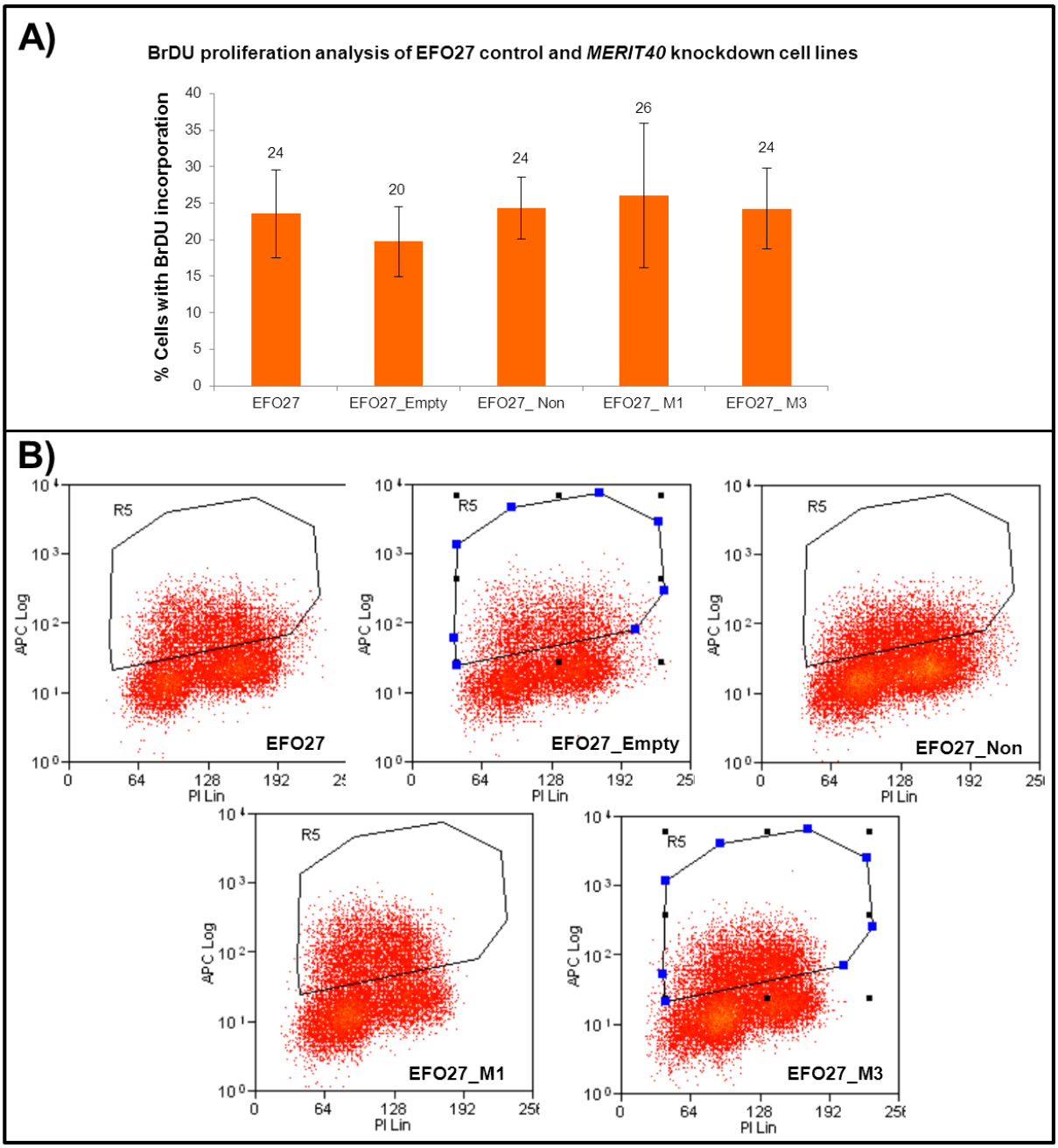


Figure 6.16**:** FACS proliferation assay of EFO27 control cell lines and *MERIT40* knockdown cell lines using BrDU. BrDU incorporates in the DNA of cells in the S phase of the cell cycle. A FITC linked anti-BrDU antibody was used to label the BrDU incorporating cells. A) Histogram showing the % of cells in the S phase, B) FACS profiles where the cells in S phase are gated (gate denoted as R5). The x axis is a measure of PI and the y axis a measure of the BrDU incorporation. According to this experiment the amount of cells in the S phase corresponding to their proliferation efficiency did not change when *MERIT40* was knocked down in EFO27.

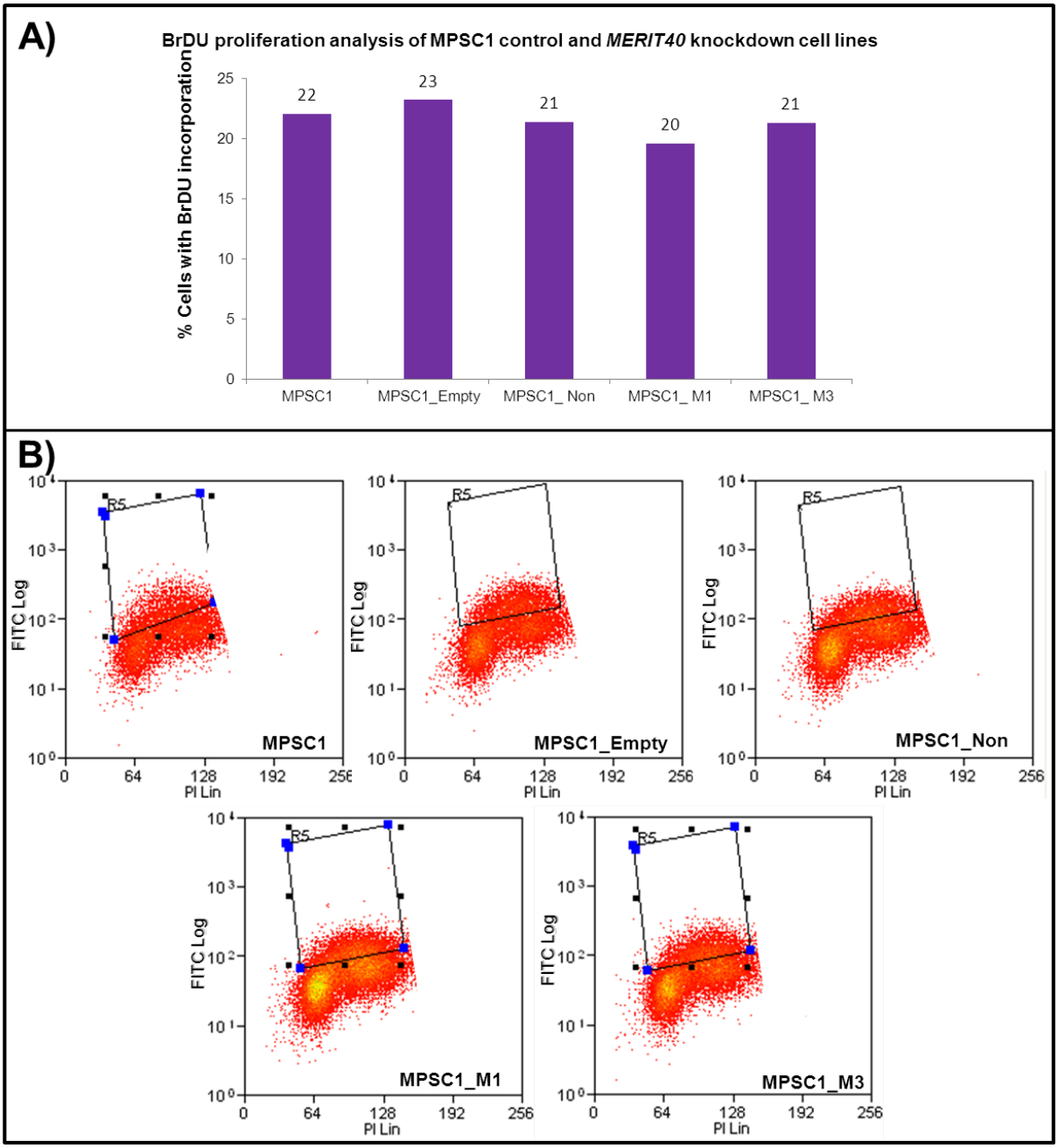


Figure 6.17**:** FACS proliferation assay of MPSC1 control cell lines and *MERIT40* knockdown cell lines using BrDU. A FITC linked anti-BrDU antibody was used to label the BrDU incorporating cells. A) Histogram showing the % of cells in the S phase, B) FACS profiles where the cells in S phase are gated (gate denoted as R5). The x axis is a measure of PI and the y axis a measure of the FITC accounting for BrDU incorporation.

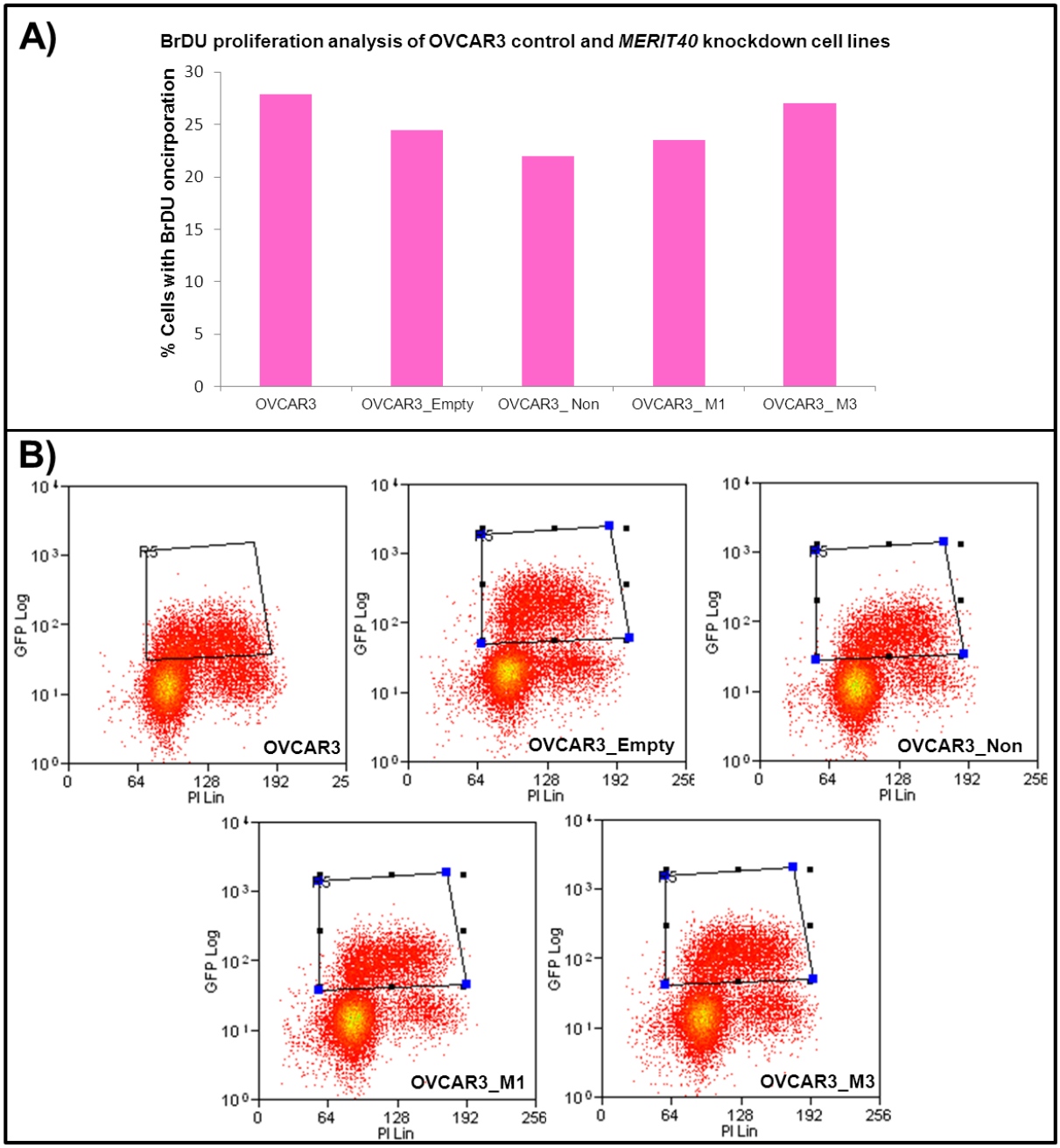


Figure 6.18**:** FACS proliferation assay of OVCAR3 control cell lines and *MERIT40* knockdown cell lines using BrDU. A FITC linked anti-BrDU antibody was used to label the BrDU incorporating cells. A) Histogram showing the % of cells in the S phase, B) FACS profiles where the cells in S phase are gated (gate denoted as R5). The x axis is a measure of PI and the y axis a measure of the FITC accounting for BrDU incorporation.

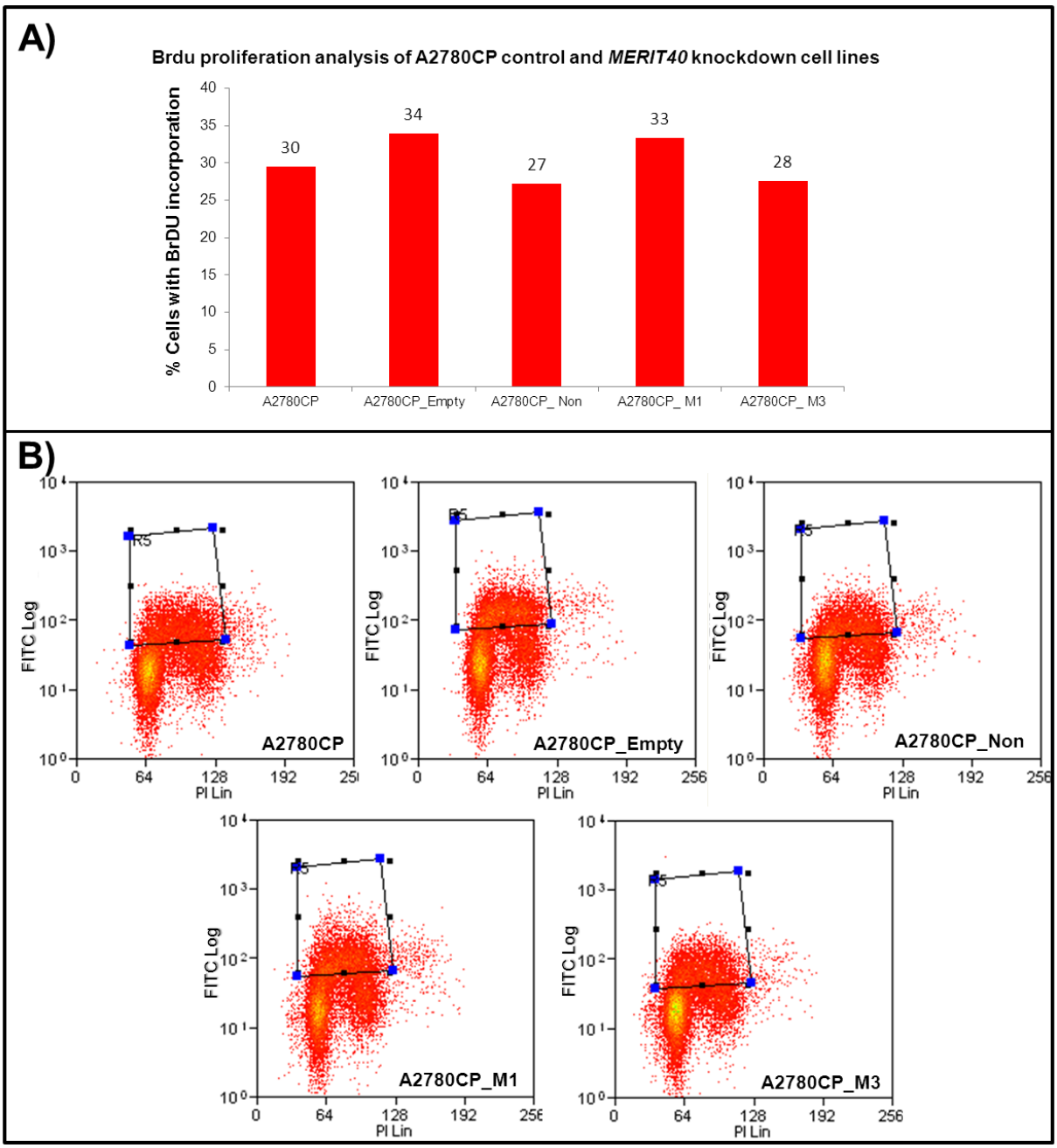


Figure 6.19**:** FACS proliferation assay of A2780CP control cell lines and *MERIT40* knockdown cell lines using BrDU. A FITC linked anti-BrDU antibody was used to label the BrDU incorporating cells. A) Histogram showing the % of cells in the S phase, B) FACS profiles where the cells in S phase are gated (gate denoted as R5). The x axis is a measure of PI and the y axis a measure of the FITC accounting for BrDU incorporation.

6.3.2 Effect of *MERIT40* knockdown on EOC cell lines on anchorage independent growth and migration efficiency of the cell lines.

A common phenotypic characteristic of neoplastic transformation is the ability of cancer cells to proliferate independently of signals that restrain growth such as contact inhibition responses. Soft agar colony formation assay is a common method to monitor anchorage independent growth and is considered an accurate *in vitro* assay for detection of malignant transformation of cells**.** I have performed soft agar colony formation assays for the EOC cell lines and their *MERIT40* knockdown derivatives to investigate whether the silencing of *MERIT40* would cause a reversal of the particular neoplastic phenotype of anchorage independent growth. The ability for anchorage independent growth was measured by the colony forming efficiency (CFE) of each cell line over a period of 4 weeks in soft agar after seeding 500, 2500 and 5000 cells in triplicate per cell line as described in the methods section. As a positive control for anchorage independent growth I used the EOC cell line TOV112D which has been previously shown to have a high colony forming efficiency in soft agar in Chapter 2.

I found that the colony forming efficiency of OVCAR3_M1 and M3 was reduced by ~7 fold compared to the average of the control cell lines (P= 0.034 and P=0.026 respectively, Figure 6.20). However it should be noted that although the number of colonies was not significantly reduced for OVCAR3_Non cell lines compared of the other controls, the size of the colonies was significantly reduced as seen on Figure 6.20. Thus, this result should be treated with caution. MPSC1 did not have the ability for anchorage independent growth as no more that 10-20 colonies in all the capture planes could be counted. None of the other two EOC cell lines showed a reduction in anchorage independent growth when *MERIT40* was knocked down. The summary of the averaged CFE of the three replicates per numbers of cells seeded per cell line is shown on Table 6.6 and the average of the CFE of all the number of cells seeded for each cell line is presented by the histograms in Figure 6.20,A. Representative pictures of the soft agars for A2780CP, MSPC1 and EFO27 are shown in Appendix 4, Figure 2.

	Colony forming efficiency (CFE)				
Number of cells seeded	**EFO27**	**EFO27_Empty**	**EFO27_Non**	**EFO27_M1**	**EFO27_M3**
5000	3	6	5	5	6
2500	3	6	5	4	6
500	4	8	6	6	6
CFE average %	**4**	**7**	**5**	**5**	**6**
	Colony forming efficiency (CFE)				
Number of cells seeded	**MPSC1**	**MPSC1_Empty**	**MPSC1_Non**	**MPSC1_M1**	**MPSC1_M3**
5000	0.33	0.35	0.36	0.49	0.47
2500	0.27	0.36	0.45	0.69	0.67
500	0.67	0.73	N/A	N/A	N/A
CFE average %	**0.42**	**0.48**	**0.41**	**0.59**	**0.57**
	Colony forming efficiency (CFE)				
Number of cells seeded	**OVCAR3**	**OVCAR3_Empty**	**OVCAR3_Non**	**OVCAR3_M1**	**OVCAR3_M3**
5000	9	6	4	0	1
2500	12	7	6	1	0
500	7	N/A	N/A	1	0
CFE average %	**9**	**7**	**5**	**0.70**	**0.64**
	Colony forming efficiency (CFE)				
Number of cells seeded	**A2780CP**	**A2780CP_Empty**	**A2780CP_Non**	**A2780CP_M1**	**A2780CP_M3**
5000	19	13	16	15	22
2500	22	20	20	11	24
500	55	41	48	27	51
CFE average %	**32**	**25**	**28**	**18**	**32**

Table 6.6: Colony forming efficiency of control and *MERIT40* knockdown EOC cell lines in soft agar assays. Summarised are the averaged CFE of the three replicates for the three numbers of cells plated for each EOC control cell lines and their *MERIT40* knockdown derivatives. N/A: not applicable: these are conditions where either there was contamination or after staining the colonies were not visible enough to be reliably counted.

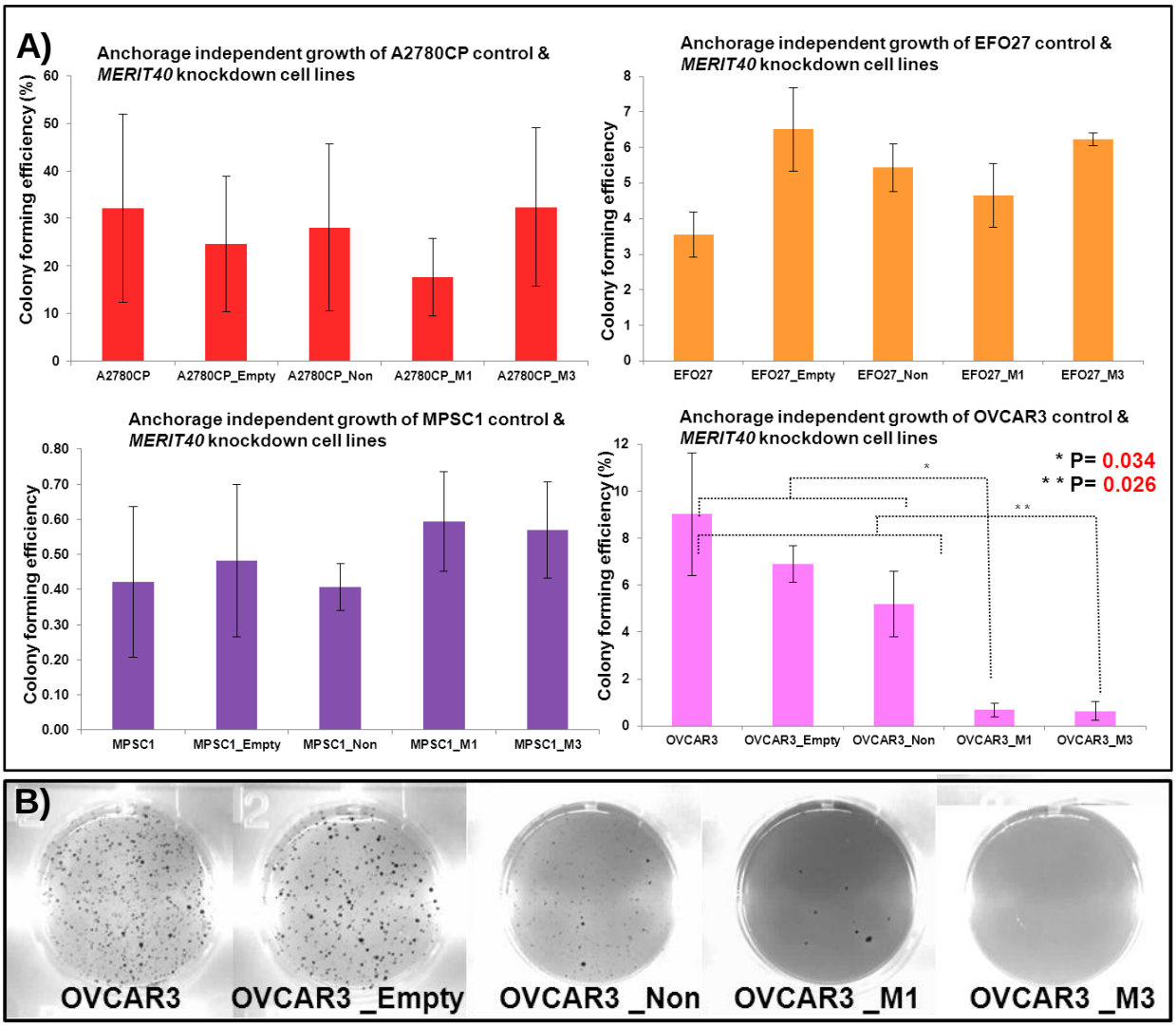


Figure 6.20**:** Anchorage independent growth assays for control and *MERIT40* knockdown EOC cell lines. A) Histograms representing the average colony forming efficiency for the control EOC cell lines compared to the *MERIT40* knockdown cell lines. There was a significant reduction observed in the anchorage independent growth ability of OVCAR3 when *MERIT40* was knocked down by both the M1 and M3 constructs. B) Representative pictures of colonies grown for the OVCAR3 controls and *MERIT40* silenced OVCAR3 cell lines after seeding of 5000 cells in soft agar over 4 weeks.

Another phenotypic characteristic of malignant cells is their increased ability to migrate. A popular way to test this is by performing trans-well assays, where the cells are plated on a membrane of a well that is positioned on top of a well filled with chemoattractant such as foetal bovine serum. The migration ability is assessed on the basis of the ability of the cells to migrate through the membrane towards the chemoattractant. I have performed the migration assay for the EOC cell line EFO27 compared to the EFO27 *MERIT40* knockdown cell lines and found no difference in the migration ability of the cells line when

MERIT40 was knocked down. The migration assay was performed over 20 hours and 3 replicates of each cell line were prepared and 5 pictures per replicate were captured, the migrated cells were counted and averaged. There was no difference observed in the ability of the cell line to migrate when *MERIT40* was knocked down. The summarised graph and representative pictures for the migration assay performed for EFO27 cell line are shown in Figure 6.21. The migration assays were not performed for the other cell lines due to the cost of the trans-wells.

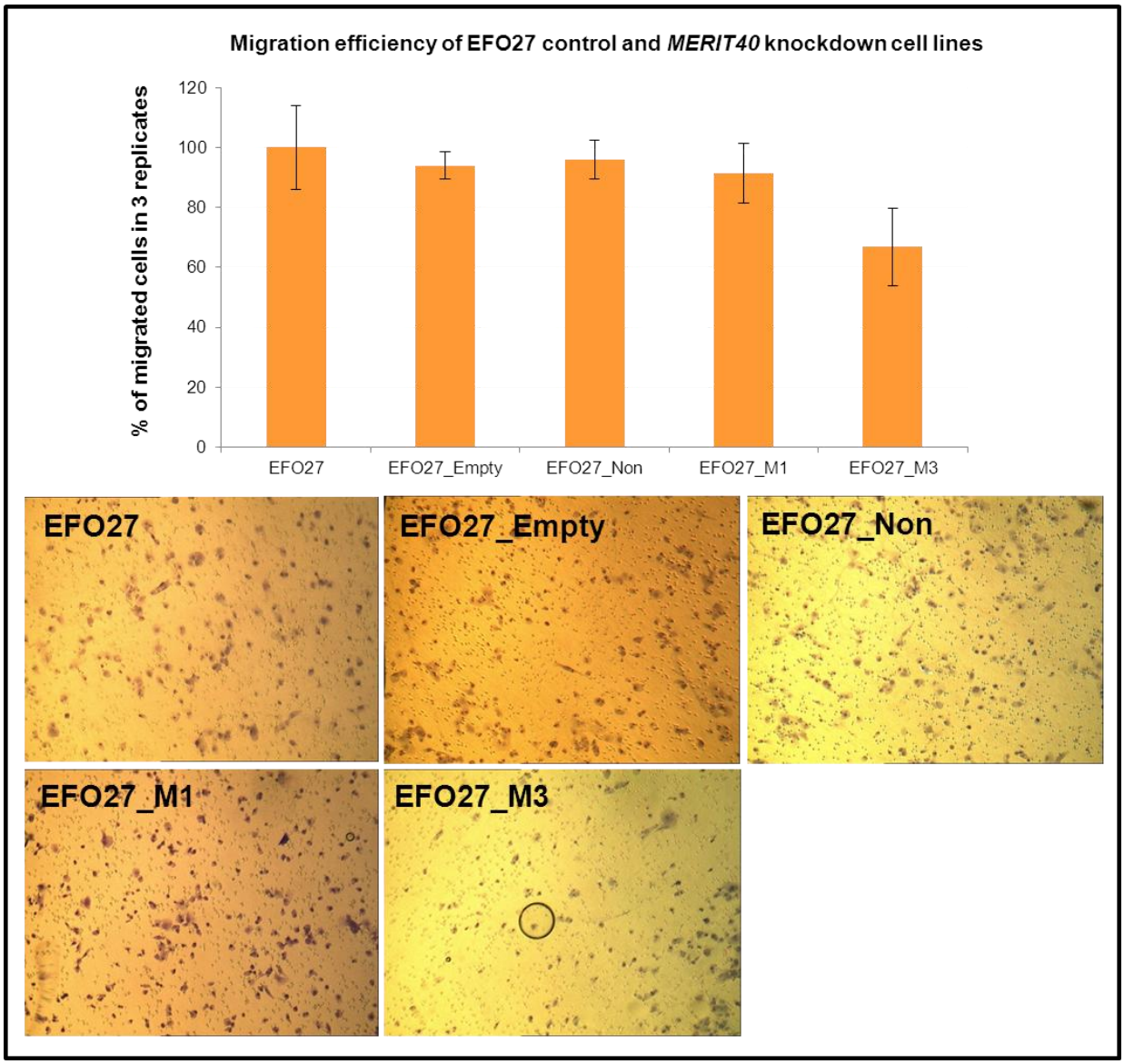


Figure 6.21**:** Migration efficiency of EFO27 control and *MERIT40* knockdown cell lines. (TOP) Histogram summarising the migration efficiency of EFO27 control and *MERIT40* knockdown cell lines. Below the histogram are shown representative pictures for the membranes of each of the EFO27 cell line.

6.3.3 Evaluating the effects of *MERIT40* knockdown on regulating other pathways involved in tumorigenesis

I have investigated in EOC cell lines with knocked down *MERIT40* the mRNA expression of 55 genes involved in pathways implicated with cell cycle progression, chromosome segregation, cell differentiation, apoptosis, DNA damage repair as well as MERIT40 interacting proteins in the BRCA1 DNA repair complex. I also investigated if the expression of other genes closely located to the *MERIT40* chromosomal locus (19p13) would be affected by the knockdown of *MERIT40*. A total of 55 genes were chosen to be investigated and the pathways they are involved are summarised on Table 6.7. The mRNA expression of the selected genes was analysed by performing a Fluidigm gene expression experiment using Taqman Real Time PCR assays. The assays were multiplexed in 7 different chips/sets with endogenous controls GAPDH and β-actin included in each chip for the mRNA expression of the selected genes to be normalised against.

The total number of samples in this experiment was 16 (Parental, Empty, Non-silencing, M1 and M3 conditions per EOC cell line MPCS1, EFO27, OVCAR3 and A2780CP). The following quality control analysis was performed. Assays with a pass rate of less than 80% were excluded from the analysis. However, the four groups of cell lines for each of the infected EOC cell lines were also analysed separately. Thus in cases that all conditions of one EOC cell line failed bringing the statistical value of the overall assay pass rate to less than 80%, the rest of the groups were analysed regardless (for example for CDKN2A expression assay all MSPC1 control and *MERIT40* knockdown cell lines, Table 6.7). Samples with <85% pass rate were also excluded. Control EOC cell line A2780CP was excluded, therefore I have used as a calibrator the control cell lines EOC_Empty for all stable *MERIT40* knockdown cell lines generated and used in this experiment. The cut-off for accepted standard deviation between the three replicates was <0.6 otherwise the outlier replicate was excluded from the analysis as long there would be two replicates remaining. In rare cases to keep two replicates in, a standard deviation between 0.6 and 1 was accepted but in assays where this occurred in more than 15% of the samples the assay was excluded. Finally, a standard curve was generated

for each of the assays and if any assay exhibited an R^2 <0.9 it was failed. The standard curve graphs for all the assays performed are shown in Appendix 4, Figure 3. After quality control analysis 8 assays were excluded and β-actin for set 2 was also excluded. A summary of the assays that were excluded from the analysis and the reasons for excluding them are shown in Table 6.8.

Gene name	Pathway involved/ Reason for selection
CCNE1	Cell cycle control, G1/S transition
CDKN1A	Cell cycle control, G1/S transition
CDKN2A	Cell cycle control/arrest in G1 and G2
PTEN	Cell cycle control
RB1	Cell cycle progression inhibitor
BUBR1	Cell cycle Mitotic checkpoint control. Kinetochore localisation
MAD1	Cell cycle Mitotic checkpoint control. Kinetochore localisation
MAD2L1	Cell cycle Mitotic checkpoint control. Kinetochore localisation
MAD2L2	Cell cycle Mitotic checkpoint control. Kinetochore localisation
CHK1	Cell cycle control, entry to mitosis
CHK2	Cell cycle control, entry to mitosis
BUB1	Chromosome segregation
BUB3	Chromosome segregation
ATM	Activates checkpoint signalling upon DNA double strand breaks (DSBs)
FANCD2	DNA damage response by HR.
PARP1	DNA damage response by HR.
RAD51	DNA damage response by HR.
RAD51C	DNA damage response by HR.
RAD51D	DNA damage response by HR.
BARD1	DNA damage response by HR
BRCA1	DNA damage response by HR.
BRCA2	DNA damage response by HR.
MERIT40	DNA damage response by HR. BRCA1 complex
BRE	DNA damage response by HR. MERIT40 interacting
CCDC98	DNA damage response by HR. MERIT40 interacting
RAP80	DNA damage response by HR. MERIT40 interacting
BRCC3	DNA damage response by HR. MERIT40 interacting
ERCC1	DNA damage response. Nucleotide excision repair
ERCC2	DNA damage response. Nucleotide excision repair
XPA	DNA damage response. Nucleotide excision repair
XRCC5	DNA damage response. Nucleotide excision repair
XRCC6	DNA damage response. Nucleotide excision repair
POLB	DNA damage response by base excision repair
MLH1	DNA damage signalling. Involved in mismatch repair
MSH2	DNA damage signalling. Involved in mismatch repair
PMS1	DNA damage signalling. Involved in mismatch repair
PIK3CA	Cell proliferation regulator
KRAS	Cell growth and proliferation
BRAF	Cell proliferation, differentiation, apoptosis. MAPK/ERK signalling

HOXA10	Cell differentiation
CDK12	Cell proliferation, cell differentiation, apoptosis
BAD	Apoptosis regulator
TP53	Apoptosis regulator
ABHD8	locus 19p13
CPAMD8	locus 19p13
DDA1	locus 19p13
GTPBP3	locus 19p13
MRPL34	locus 19p13
MYO9B	locus 19p13
NR2F6	locus 19p13
OCEL1	locus 19p13
USHBP1	locus 19p13
ANKRD41	locus 19p13
TMEM16H	locus 19p13
USE1	locus 19p13

Table 6.7**:** List of the genes their expression to be investigated when *MERIT40* was knocked down in EOC cell lines. Summarised in the table is a brief summary of the genes' function. The function of genes at locus 19p13 have been described in chapter 5.

Gene	**Fluidigm Chip**	**Assay pass rate %**	**% of samples with >0.6SD**	**R^2 Standard Curve**	**Assay passed or failed**
CCNE1	Chip 1	100	0	0.991	PASSED
CDKN1A	Chip 5	100	0	0.992	PASSED
CDKN2A	Chip 5	75 *	0	0.998	PASSED
PTEN	Chip 5	100	0	0.954	PASSED
RB1	Chip 7	100	0	0.997	PASSED
BUBR1	Chip 4	100	0	0.967	PASSED
MAD1	Chip 5	100	6.25	0.921	PASSED
MAD2L1	Chip 5	100	0	0.982	PASSED
MAD2L2	Chip 5	100	0	0.986	PASSED
CHK1	Chip 5	100	0	0.993	PASSED
CHK2	Chip 5	75 *	8.3	0.997	PASSED
BUB1	Chip 4	100	0	0.965	PASSED
BUB3	Chip 4	100	0	0.998	PASSED
ATM	Chip 4	100	0	0.981	PASSED
FANCD2	Chip 5	100	0	0.947	PASSED
PARP1	Chip 5	100	0	0.997	PASSED
RAD51	Chip 5	100	0	0.946	PASSED
RAD51C	Chip 5	100	0	0.969	PASSED
RAD51D	Chip 5	100	0	0.971	PASSED
BARD1	Chip 5	100	0	0.942	PASSED
BRCA1	Chip 4	100	0	0.997	PASSED
BRCA2	Chip 4	100	0	0.998	PASSED
MERIT40	Chip 2	100	0	0.960	PASSED
BRE	Chip 4	100	0	0.997	PASSED

CCDC98	Chip 4	100	0	0.982	PASSED
RAP80	Chip 5	100	0	0.987	PASSED
BRCC3	Chip 1	100	0	0.992	PASSED
ERCC1	Chip 5	100	0	0.998	PASSED
ERCC2	Chip 5	100	0	0.977	PASSED
XPA	Chip 5	75 **	31.25	0.626	FAILED
XRCC5	Chip 5	100	0	0.997	PASSED
XRCC6	Chip 5	100	0	0.989	PASSED
POLB	Chip 5	100	0	0.939	PASSED
MLH1	Chip 5	50 ***	0	0.976	PASSED
MSH2	Chip 5	100	0	0.993	PASSED
PMS1	Chip 5	100	12.5	0.932	PASSED
PIK3CA	Chip 5	100	0	0.984	PASSED
KRAS	Chip 5	100	0	0.972	PASSED
BRAF	Chip 4	100	0	0.967	PASSED
HOXA10	Chip 5	50 ***	0	0.488	FAILED
CDK12	Chip 6	100	0	0.996	PASSED
BAD	Chip 4	50	31.5	0.585	FAILED
TP53	Chip 5	100	0	0.995	PASSED
ABHD8	Chip 2	100	0	0.922	PASSED
CPAMD8	Chip 2	0	N/A	0.653	FAILED
DDA1	Chip 2	100	0	0.938	PASSED
GTPBP3	Chip 2	100	0	0.925	PASSED
MRPL34	Chip 3	100	0	0.972	PASSED
MYO9B	Chip 3	100	0	0.983	PASSED
NR2F6	Chip 3	100	0	0.944	PASSED
OCEL1	Chip 3	100	0	0.890	FAILED
USHBP1	Chip 4	68.8	50	FAILED	FAILED
ANKRD41	Chip 4	100	6.25	0.717	FAILED
TMEM16H	Chip 4	100	12.5	0.317	FAILED
USE1	Chip 4	100	0	0.906	PASSED
β-actin SET 1	Chip 1	100	0	0.998	PASSED
GAPDH SET 1	Chip 1	100	0	0.996	PASSED
β-actin SET 2	Chip 2	100	6.25	0.633	FAILED
GAPDH SET 2	Chip 2	100	0	0.997	PASSED
β-actin SET 3	Chip 3	100	0	0.979	PASSED
GAPDH SET 3	Chip 3	100	0	0.994	PASSED
β-actin SET 4	Chip 4	100	0	0.996	PASSED
GAPDH SET 4	Chip 4	100	0	0.999	PASSED
β-actin SET 5	Chip 5	100	0	0.994	PASSED
GAPDH SET 5	Chip 5	100	0	0.996	PASSED
β-actin SET 6	Chip 6	100	0	0.997	PASSED
GAPDH SET 6	Chip 6	100	0	0.997	PASSED
GAPDH SET 7	Chip 7	100	0	0.995	PASSED

Table 6.8**:** Quality control analysis for Fluidigm gene expression experiment for the control EOC and *MERIT40* knockdown cell lines. Highlighted in red are the failed assays for a particular quality control criterion. (*) only failed for OVCAR3 control and *MERIT40* knockdown cell lines. (**) only failed for MPSC1 control and *MERIT40* knockdown cell lines. (***) only failed for OVCAR3 and A2780CP control and *MERIT40* knockdown cell lines.

The expression of the selected genes in EOC cell lines with knocked down *MERIT40* was then calculated using the relative ΔΔCt method and calibrated against the expression of the respective EOC cell lines infected with a virus containing the Empty vector normalised against the appropriate endogenous control(s). For a difference in expression to be biologically meaningful at least a two-fold increase or decrease would have to be observed. The fold change was calculated as the ratio of the expression in the knockdown versus the expression of the EOC_Empty calibrator sample. A fold change of <0.5 indicates a two-fold under-expression of the gene and a fold change of >2 indicated a two-fold over-expression of the gene. Although the EOC_Empty cell lines were used as the calibrator sample, the EOC_Non samples were also serving as controls for additional quality control. The fold change between the EOC_Empty and the EOC_Non samples was also calculated and if was found to have be <0.5 or >2 the assay was deemed as having a fold change discordance and any observed fold changes observed between the controls and knockdown samples were not considered reliable. Similarly, if a biologically significant fold change was observed between EOC_Empty and knockdown cell lines and was not observed between EOC_Non and the respective knockdown cell line it was classed as having a fold change discordance and the positive result was also disregarded. Finally, it is possible that a biologically significant observed fold change would not be statistically significant. Fold change calculations are calculated using the mean relative expressions and do not take into account the diversity in the expression of the replicates. Thus, in cases were significant fold changes were observed, a two tailed paired student t-test was performed further to judge on whether the observed fold change would also be statistically significant ($P<0.05$).

The *MERIT40* knockdown efficiency of the different shRNA hairpins observed previously (Figure 6.6) was confirmed in the present Fluidigm gene expression experiment, further indicating the robustness of the assay. For EFO27 both M1 and M3 showed a sufficient 9-15 and 2-3 fold decrease respectively in *MERIT40* expression confirming the 14 and 2.5 fold decrease observed in the Taqman gene expression assay (Figure 6.6). For MPSC1 and A2780CP only M1 achieved a sufficient decrease of *MERIT40* expression of 8-9

and 2-2.5 fold decrease respectively which was similar to the calculated 8 and 3 fold decrease of observed in the Taqman gene expression assay. For OVCAR3 M1 showed a 7-9 fold decrease of *MERIT40* expression similar to the 7 fold decrease found by the Taqman gene expression assay. Also confirming the previous observations, M3 did not achieve a more than two-fold decrease of the gene's expression. Figure 6.22 presents the differential expression of *MERIT40* in the EOC and *MERIT40* knockdown cell lines according to the Fluidigm gene expression assay. The biologically significant fold change in the expression of *MERIT40* was also strengthened by statistical significance after performing a student's t-test (Figure 6.22).

The differential expression of the selected genes between the control and *MERIT40* knockdown cell lines was evaluated. A summary of the expression values for all the genes in the different cell lines is presented in Appendix 4, Tables 1, 2, 3 and 4. The results are presented in the 4 different cell lines separately as they are representing difference subtypes of ovarian cancer and they showed different behaviours in some of the phenotypic assays previously performed. An interesting gene to evaluate was *PARP1.* PARP1 pathway could be an alternative repair mechanism the cells employ when MERIT40 assisted HR is defected to defend themselves against DNA damage caused by platinum drugs. No differences were observed in the *PARP1* mRNA expression for any of the 4 cell lines studied between the controls and their *MERIT40* knockdown counterparts.

For EFO27, none of the genes was differentially expressed in the presence or absence of MERIT40 (Table 6.9). I have previously observed a reduction of cells in the G2/M phase and reduction in the cells with 4N content for the EFO27_M1 and M3 compared to the parental. This observation could not be linked to any differences in the expression of genes involved in the cell cycle regulation or chromosome segregation when *MERIT40* was knocked down in the EFO27 EOC cell line. The fold changes of all the selected genes between the control and *MERIT40* knockdown EFO27 cell lines are presented in Table 6.9.

For MSPC1 cell line two genes were found to be under-expressed when *MERIT40* was knocked down by the M1 construct. One of those genes was

BRE, coding for a component of the BRCA1 complex which is required for the integrity of the BRCA1 complex and localisation of the site of DNA DSBs. *BRE* was found to be 2 fold under-expressed in the MSCP1_M1 cell line compared to the controls. The result was not of statistical significance though with P values>0.05 (Figure 6.23). The second gene found with a 2 fold under-expression in the MPSC1_M1 and M3 cell lines compared to the controls was *BRAF,* which is involved in cell proliferation, with a statistically significant P value of <0.05 (Figure 6.23). The fold changes of all the selected genes between the control and *MERIT40* knockdown MPSC1 cell lines are presented in Table 6.10.

For OVCAR3 cell line two genes were found to be under-expressed when *MERIT40* was knocked down by the M1 construct. One of those genes was *RAD51D*, coding for a protein involved in HR and is part of the RAD51 complex responsible for searching the genome for an intact copy of the damaged DNA on the sister chromatid. *RAD51D* was found to be 2 fold under-expressed (with P<0.05) in the OVCAR3 _M1 cell line compared to the controls. The second gene with a 2.5 fold under-expression in the OVCAR3 _M1 cell line was *ABHD8,* which is a gene located in chromosome 19 in close proximity to the *MERIT40* locus. ABHD8 may have catalytic and hydrolase activity and might be involved in metabolic processes. Another gene was found to be 2-3 fold over-expressed (with P<0.05) when *MERIT40* was knocked down in OVCAR3 cell line was *GTPBP3*, which is another gene in close proximity to *MERIT40* in chromosome 19. According to the QC criteria this result should be taken with caution as a 2 fold change was observed between OVCAR3_Empty and OVCAR3_Non but with a trend of under-expression for OVCAR_Non compared to OVCAR3_Empty. However, the expression difference observed between the OVCAR3_Empty and OVCAR3_Non was not statistically significant with a P value of >0.05 and the expression differences between the control and knockdown cell lines were statistically significant (Figure 6.23). The fold changes of all the selected genes between the control and *MERIT40* knockdown OVCAR3 cell lines are presented in Table 6.11.

For cell line A2780CP, *BARD1*, a gene involved in DDR in the G1/S checkpoint in response to ionising radiation (Roy et al, 2012), was found to be

2.5 fold over-expressed when *MERIT40* was knocked down (P< 0.05, Figure 6.23). The fold changes of all the selected genes between the control and *MERIT40* knockdown A2780CP cell lines are presented in Table 6.12.

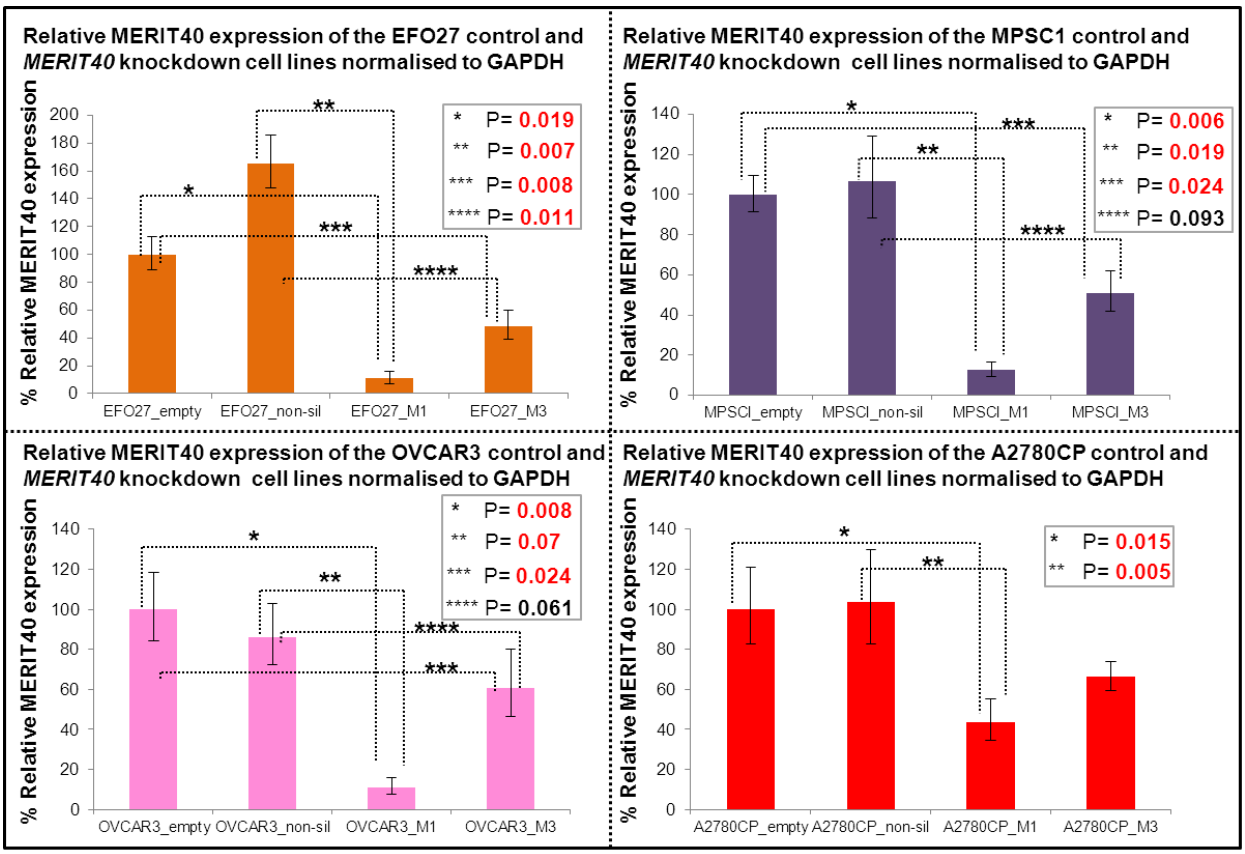


Figure 6.22**:** Validation of differential expression of *MERIT40* between the EOC control and *MERIT40* knockdown cell lines in the Fluidigm experiment. The P values of <0.05 show the statistically significant differential expression observed between the control and knockdown cell lines.

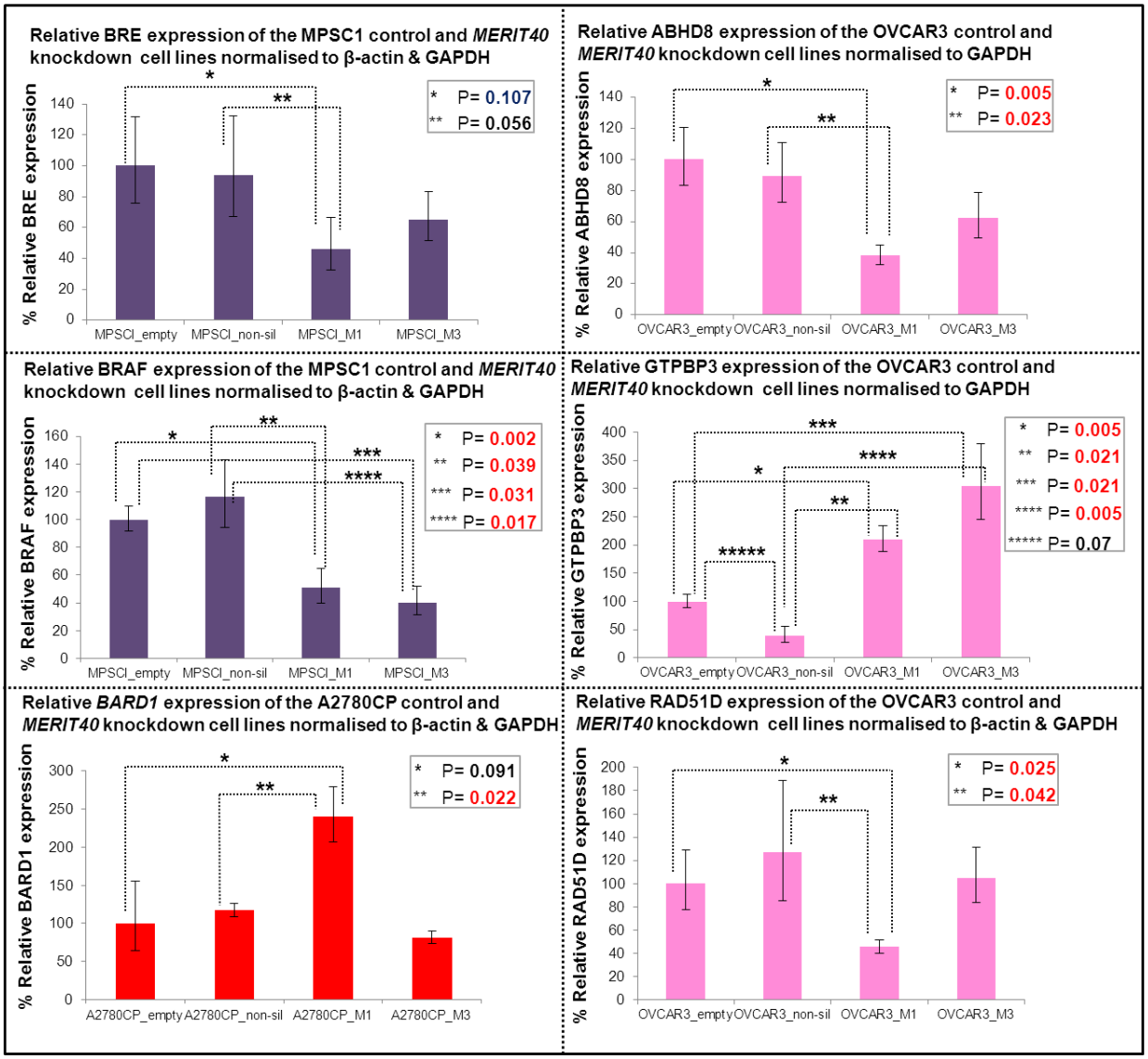


Figure 6.23**:** Histograms representing the expression of genes that were found to be differentially expressed between control and *MERIT40* knockdown cell lines. The histograms in purple are showing differential expression observed for *BRE* and *BRAF* genes between the MPSC1 control and knockdown cell lines. In pink differential expression observed for *ABHD8*, *GTPBP3* and *RAD51D* genes between the OVCAR3 control and knockdown cell lines. In red showing the differential expression observed for *BARD1* between the A2780CP control and knockdown cell lines. For those genes that showed a significant fold difference, the P values were calculated to check whether the change would be statistically significant and annotated by the graphs. All fold changes observed were statistically significant apart for the one observed for the gene *BRE* in the MPSC1 cell line.

Gene	Fluidigm Chip	Normalised against	Fold Change					Fold change concordance
			Non vs E	M1 vs E	M3 vs E	M1 vs Non	M3 vs Non	
CCNE1	Chip 1	β-actin & GAPDH	1.9	1.85	2.03	0.97	1.07	NO
CDKN1A	Chip 5	β-actin & GAPDH	1.21	0.65	1.12	0.54	0.93	YES
CDKN2A	Chip 5	β-actin & GAPDH	1.4	1.06	1.51	0.76	1.08	YES
PTEN	Chip 5	β-actin & GAPDH	0.93	1.04	1.17	1.12	1.26	YES
RB1	Chip 7	β-actin & GAPDH	1.01	0.85	0.8	0.84	0.79	YES
BUBR1	Chip 4	β-actin & GAPDH	1.03	0.97	0.95	0.94	0.92	YES
MAD1	Chip 5	β-actin & GAPDH	0.99	0.57	0.69	0.58	0.70	YES
MAD2L1	Chip 5	β-actin & GAPDH	1.6	0.92	1.14	0.58	0.71	YES
MAD2L2	Chip 5	β-actin & GAPDH	1.19	1.16	1.2	0.97	1.01	YES
CHK1	Chip 5	β-actin & GAPDH	1.22	1.12	1.12	0.92	0.92	YES
CHK2	Chip 5	β-actin & GAPDH	1.53	0.76	1.54	0.50	1.01	NO
BUB1	Chip 4	β-actin & GAPDH	1.22	1.04	1.22	0.85	1.00	YES
BUB3	Chip 4	β-actin & GAPDH	0.96	0.97	1.1	1.01	1.15	YES
ATM	Chip 4	β-actin & GAPDH	0.8	0.99	0.88	1.24	1.10	YES
FANCD2	Chip 5	β-actin & GAPDH	1.5	1.54	1.23	1.03	0.82	YES
PARP1	Chip 5	β-actin & GAPDH	1.57	1.23	1.5	0.78	0.96	YES
RAD51	Chip 5	β-actin & GAPDH	1.74	1.35	1.1	0.78	0.63	YES
RAD51C	Chip 5	β-actin & GAPDH	1.94	1.38	1.27	0.71	0.65	YES
RAD51D	Chip 5	β-actin & GAPDH	1.57	1.66	1.5	1.06	0.96	YES
BARD1	Chip 5	β-actin & GAPDH	0.83	1.03	1.9	1.24	2.29	NO
BRCA1	Chip 4	β-actin & GAPDH	1.06	0.92	1.01	0.87	0.95	YES
BRCA2	Chip 4	β-actin & GAPDH	0.89	1.04	1	1.17	1.12	YES
MERIT40	Chip 2	β-actin	1.65	0.11	0.49	0.07	0.30	YES
BRE	Chip 4	β-actin & GAPDH	0.76	0.89	0.98	1.17	1.29	YES
CCDC98	Chip 4	β-actin & GAPDH	0.97	0.6	0.96	0.62	0.99	YES
RAP80	Chip 5	β-actin & GAPDH	1.68	1.22	1.36	0.73	0.81	YES
BRCC3	Chip 1	β-actin & GAPDH	1.56	1.05	1.29	0.67	0.83	YES
ERCC1	Chip 5	β-actin & GAPDH	1.09	0.94	1.32	0.86	1.21	YES
ERCC2	Chip 5	β-actin & GAPDH	1.43	1.39	1.29	0.97	0.90	YES
XRCC5	Chip 5	β-actin & GAPDH	1.33	0.94	1.25	0.71	0.94	YES
XRCC6	Chip 5	β-actin & GAPDH	1.32	1.26	1.2	0.95	0.91	YES
POLB	Chip 5	β-actin & GAPDH	1.81	2.34	1.64	1.29	0.91	NO
MLH1	Chip 5	β-actin & GAPDH	1.81	1.34	1.55	0.74	0.86	YES
MSH2	Chip 5	β-actin & GAPDH	1.38	1.25	1.22	0.91	0.88	YES
PMS1	Chip 5	β-actin & GAPDH	2.82	2.89	1.18	1.02	0.42	NO
PIK3CA	Chip 5	β-actin & GAPDH	1.43	1.32	0.96	0.92	0.67	YES
KRAS	Chip 5	β-actin & GAPDH	1.53	1.33	1.22	0.87	0.80	YES
BRAF	Chip 4	β-actin & GAPDH	0.6	1.08	1.12	1.80	1.87	YES
CDK12	Chip 6	β-actin & GAPDH	0.76	0.54	0.72	0.71	0.95	YES
TP53	Chip 5	β-actin & GAPDH	1.04	1	0.96	0.96	0.92	YES
ABHD8	Chip 2	β-actin	1.31	0.92	1.06	0.70	0.81	YES
DDA1	Chip 2	β-actin	1.11	1.12	1.27	1.01	1.14	YES
GTPBP3	Chip 2	β-actin	0.82	1.26	1.81	1.54	2.21	NO
MRPL34	Chip 3	β-actin & GAPDH	0.95	1.11	1.4	1.17	1.47	YES
MYO9B	Chip 3	β-actin & GAPDH	1.18	0.97	0.95	0.82	0.81	YES
NR2F6	Chip 3	β-actin & GAPDH	0.65	0.85	1.39	1.31	2.14	NO
USE1	Chip 4	β-actin & GAPDH	1.23	1.63	1.83	1.33	1.49	YES

Table 6.9**:** Summary of fold changes for the selected genes in the EFO27 and silenced *MERIT40* counterpart cell lines. Highlighted in green are 2 fold changes or higher. Higher than 2 fold changes indicated over-expression of the gene and lower than 0.5 fold changed indicate under-expression of the genes. The column on the far right namely fold change concordance is indicating whether the fold changes observed are concordant or not.

Gene	Fluidigm Chip	Normalised against	Fold Change					Fold change concordance
			Non vs E	M1 vs E	M3 vs E	M1 vs Non	M3 vs Non	
CCNE1	Chip 1	β-actin & GAPDH	0.88	0.73	0.94	0.83	1.07	YES
CDKN1A	Chip 5	β-actin & GAPDH	0.97	0.86	0.96	0.89	0.99	YES
CDKN2A	Chip 5	β-actin & GAPDH	N/A	N/A	N/A	N/A	N/A	YES
PTEN	Chip 5	β-actin & GAPDH	0.86	0.81	0.92	0.94	1.07	YES
RB1	Chip 7	β-actin & GAPDH	0.99	1.01	0.99	1.02	1.00	YES
BUBR1	Chip 4	β-actin & GAPDH	1.09	0.64	1.27	0.59	1.17	YES
MAD1	Chip 5	β-actin & GAPDH	0.91	0.95	1.24	1.04	1.36	YES
MAD2L1	Chip 5	β-actin & GAPDH	1.03	0.83	1.23	0.81	1.19	YES
MAD2L2	Chip 5	β-actin & GAPDH	1.05	0.93	1.34	0.89	1.28	YES
CHK1	Chip 5	β-actin & GAPDH	1.22	0.96	1.52	0.79	1.25	YES
CHK2	Chip 5	β-actin & GAPDH	1.77	1.11	1.46	0.63	0.82	YES
BUB1	Chip 4	β-actin & GAPDH	1.26	0.58	1.24	0.46	0.98	NO
BUB3	Chip 4	β-actin & GAPDH	1.3	0.72	1.17	0.55	0.90	YES
ATM	Chip 4	β-actin & GAPDH	1.52	0.93	1.09	0.61	0.72	YES
FANCD2	Chip 5	β-actin & GAPDH	0.93	1.03	1.29	1.11	1.39	YES
PARP1	Chip 5	β-actin & GAPDH	1.01	0.96	1.42	0.95	1.41	YES
RAD51	Chip 5	β-actin & GAPDH	1.13	0.96	1.01	0.85	0.89	YES
RAD51C	Chip 5	β-actin & GAPDH	0.69	0.86	1.67	1.25	2.42	NO
RAD51D	Chip 5	β-actin & GAPDH	0.78	0.72	0.88	0.92	1.13	YES
BARD1	Chip 5	β-actin & GAPDH	0.76	0.86	0.89	1.13	1.17	YES
BRCA1	Chip 4	β-actin & GAPDH	1.17	0.62	1.04	0.53	0.89	YES
BRCA2	Chip 4	β-actin & GAPDH	1.35	0.89	1	0.66	0.74	YES
MERIT40	Chip 2	β-actin	1.07	0.12	0.51	0.11	0.48	YES
BRE	Chip 4	β-actin & GAPDH	0.94	0.46	0.65	0.49	0.69	YES
CCDC98	Chip 4	β-actin & GAPDH	1.07	0.7	0.93	0.65	0.87	YES
RAP80	Chip 5	β-actin & GAPDH	1.3	0.96	1.68	0.74	1.29	YES
BRCC3	Chip 1	β-actin & GAPDH	1.43	1	0.99	0.70	0.69	YES
ERCC1	Chip 5	β-actin & GAPDH	0.81	0.71	0.66	0.88	0.81	YES
ERCC2	Chip 5	β-actin & GAPDH	0.98	1.01	1.25	1.03	1.28	YES
XRCC5	Chip 5	β-actin & GAPDH	1.07	0.96	1.41	0.90	1.32	YES
XRCC6	Chip 5	β-actin & GAPDH	1.1	1.07	1.57	0.97	1.43	YES
POLB	Chip 5	β-actin & GAPDH	0.95	0.83	1.24	0.87	1.31	YES
MLH1	Chip 5	β-actin & GAPDH	1.01	1.03	1.03	1.02	1.02	YES
MSH2	Chip 5	β-actin & GAPDH	1.03	0.84	1.49	0.82	1.45	YES
PMS1	Chip 5	β-actin & GAPDH	0.9	1.38	1.44	1.53	1.60	YES
PIK3CA	Chip 5	β-actin & GAPDH	0.99	0.73	0.91	0.74	0.92	YES
KRAS	Chip 5	β-actin & GAPDH	0.98	0.79	0.93	0.81	0.95	YES
BRAF	Chip 4	β-actin & GAPDH	1.16	0.51	0.4	0.44	0.34	YES
CDK12	Chip 6	β-actin & GAPDH	1.16	1.06	1.06	0.91	0.91	YES
TP53	Chip 5	β-actin & GAPDH	1.09	1.08	1.06	0.99	0.97	YES
ABHD8	Chip 2	β-actin	1.24	0.83	0.55	0.67	0.44	NO
DDA1	Chip 2	β-actin	1.25	0.91	0.6	0.73	0.48	NO
GTPBP3	Chip 2	β-actin	1.28	0.95	0.79	0.74	0.62	YES
MRPL34	Chip 3	β-actin & GAPDH	3.11	1.15	0.71	0.37	0.23	NO
MYO9B	Chip 3	β-actin & GAPDH	1.29	1.15	0.78	0.89	0.60	YES
NR2F6	Chip 3	β-actin & GAPDH	1.9	1.08	0.78	0.57	0.41	NO
USE1	Chip 4	β-actin & GAPDH	0.88	0.45	0.88	0.51	1.00	NO

Table 6.10**:** Summary of the fold changes for the selected genes in the MPSC1 and silenced *MERIT40* counterpart cell lines. Highlighted in green are 2 fold changes or more. Higher than 2 fold changes indicated over-expression of the gene and lower than 0.5 fold changed indicate under-expression of the genes. The column on the far right namely fold change concordance is indicating whether the fold changes observed are concordant or not.

Gene	Fluidigm Chip	Normalised against	Fold Change					Fold change concordance
			Non vs E	M1 vs E	M3 vs E	M1 vs Non	M3 vs Non	
CCNE1	Chip 1	β-actin & GAPDH	0.89	1.14	0.65	1.28	0.73	YES
CDKN1A	Chip 5	β-actin & GAPDH	0.87	0.65	0.5	0.75	0.57	YES
CDKN2A	Chip 5	β-actin & GAPDH	0.74	0.47	0.58	0.64	0.78	NO
PTEN	Chip 5	β-actin & GAPDH	0.93	0.63	1	0.68	1.08	YES
RB1	Chip 7	β-actin & GAPDH	0.89	0.78	0.79	0.88	0.89	YES
BUBR1	Chip 4	β-actin & GAPDH	0.93	1.14	1.55	1.23	1.67	YES
MAD1	Chip 5	β-actin & GAPDH	1.33	0.74	1.03	0.56	0.77	YES
MAD2L1	Chip 5	β-actin & GAPDH	0.86	0.89	1.01	1.03	1.17	YES
MAD2L2	Chip 5	β-actin & GAPDH	0.8	0.75	0.84	0.94	1.05	YES
CHK1	Chip 5	β-actin & GAPDH	1	0.82	1.11	0.82	1.11	YES
CHK2	Chip 5	β-actin & GAPDH	N/A	N/A	N/A	N/A	N/A	N/A
BUB1	Chip 4	β-actin & GAPDH	0.75	0.68	1.05	0.91	1.40	YES
BUB3	Chip 4	β-actin & GAPDH	0.92	0.89	1.36	0.97	1.48	YES
ATM	Chip 4	β-actin & GAPDH	1.26	1.02	1.52	0.81	1.21	YES
FANCD2	Chip 5	β-actin & GAPDH	1	0.5	0.79	0.50	0.79	YES
PARP1	Chip 5	β-actin & GAPDH	1.15	0.61	0.82	0.53	0.71	YES
RAD51	Chip 5	β-actin & GAPDH	0.98	0.99	0.84	1.01	0.86	YES
RAD51C	Chip 5	β-actin & GAPDH	0.78	0.92	1.2	1.18	1.54	YES
RAD51D	Chip 5	β-actin & GAPDH	1.27	0.45	1.05	0.35	0.83	YES
BARD1	Chip 5	β-actin & GAPDH	0.67	1.03	0.77	1.54	1.15	YES
BRCA1	Chip 4	β-actin & GAPDH	0.96	1.18	1.63	1.23	1.70	YES
BRCA2	Chip 4	β-actin & GAPDH	0.86	0.99	1.09	1.15	1.27	YES
MERIT40	Chip 2	β-actin	0.86	0.11	0.61	0.13	0.71	YES
BRE	Chip 4	β-actin & GAPDH	0.87	0.8	1.53	0.92	1.76	YES
CCDC98	Chip 4	β-actin & GAPDH	0.92	0.49	1.26	0.53	1.37	NO
RAP80	Chip 5	β-actin & GAPDH	0.82	0.63	0.77	0.77	0.94	YES
BRCC3	Chip 1	β-actin & GAPDH	0.84	0.95	0.94	1.13	1.12	YES
ERCC1	Chip 5	β-actin & GAPDH	0.55	0.41	0.55	0.75	1.00	NO
ERCC2	Chip 5	β-actin & GAPDH	1.17	0.75	0.94	0.64	0.80	YES
XRCC5	Chip 5	β-actin & GAPDH	0.8	0.62	0.96	0.78	1.20	YES
XRCC6	Chip 5	β-actin & GAPDH	1.25	0.7	1.04	0.56	0.83	YES
POLB	Chip 5	β-actin & GAPDH	1.25	0.77	0.97	0.62	0.78	YES
MLH1	Chip 5	β-actin & GAPDH	N/A	N/A	N/A	N/A	N/A	N/A
MSH2	Chip 5	β-actin & GAPDH	1.44	0.91	1.13	0.63	0.78	YES
PMS1	Chip 5	β-actin & GAPDH	0.34	0.59	0.76	1.74	2.24	NO
PIK3CA	Chip 5	β-actin & GAPDH	0.91	0.56	0.78	0.62	0.86	YES
KRAS	Chip 5	β-actin & GAPDH	0.84	0.62	0.75	0.74	0.89	YES
BRAF	Chip 4	β-actin & GAPDH	1.09	1.45	1.25	1.33	1.15	YES
CDK12	Chip 6	β-actin & GAPDH	1.07	0.83	1.09	0.78	1.02	YES
TP53	Chip 5	β-actin & GAPDH	0.94	0.62	0.73	0.66	0.78	YES
ABHD8	Chip 2	β-actin	0.89	0.38	0.62	0.43	0.70	YES
DDA1	Chip 2	β-actin	0.73	1.22	1.55	1.67	2.12	NO
GTPBP3	Chip 2	β-actin	0.4	2.09	3.05	5.23	7.63	NO
MRPL34	Chip 3	β-actin & GAPDH	1.03	1.96	4.26	1.90	4.14	YES
MYO9B	Chip 3	β-actin & GAPDH	1.05	0.82	1.02	0.78	0.97	YES
NR2F6	Chip 3	β-actin & GAPDH	0.79	1.99	3.48	2.52	4.41	YES
USE1	Chip 4	β-actin & GAPDH	0.65	1.05	0.9	1.62	1.38	YES

Table 6.11**:** Summary of the fold changes for the selected genes in the OVCAR3 and silenced *MERIT40* counterpart cell lines. Highlighted in green are 2 fold changes of more. Higher than 2 fold changes indicated over-expression of the gene and lower than 0.5 fold changed indicate under-expression of the genes. The column on the far right namely fold change concordance is indicating whether the fold changes observed are concordant or not. If not concordant they are highlighted in red.

Gene	Fluidigm Chip	Normalised against	Fold Change					Fold change concordance
			Non vs E	M1 vs E	M3 vs E	M1 vs Non	M3 vs Non	
CCNE1	Chip 1	β-actin & GAPDH	0.97	1.01	1.3	1.04	1.34	YES
CDKN1A	Chip 5	β-actin & GAPDH	0.19	1.23	0.41	6.47	2.16	NO
CDKN2A	Chip 5	β-actin & GAPDH	0.76	0.97	0.58	1.28	0.76	YES
PTEN	Chip 5	β-actin & GAPDH	0.66	0.83	0.71	1.26	1.08	YES
RB1	Chip 7	β-actin & GAPDH	0.8	1.02	1.15	1.28	1.44	YES
BUBR1	Chip 4	β-actin & GAPDH	1.06	0.76	0.77	0.72	0.73	YES
MAD1	Chip 5	β-actin & GAPDH	0.54	1.22	0.68	2.26	1.26	NO
MAD2L1	Chip 5	β-actin & GAPDH	0.97	1.26	1.15	1.30	1.19	YES
MAD2L2	Chip 5	β-actin & GAPDH	0.61	0.93	0.57	1.52	0.93	YES
CHK1	Chip 5	β-actin & GAPDH	0.78	0.72	0.88	0.92	1.13	YES
CHK2	Chip 5	β-actin & GAPDH	0.82	0.93	1.13	1.13	1.38	YES
BUB1	Chip 4	β-actin & GAPDH	0.82	0.59	0.76	0.72	0.93	YES
BUB3	Chip 4	β-actin & GAPDH	0.79	0.71	0.59	0.90	0.75	YES
ATM	Chip 4	β-actin & GAPDH	1.15	0.87	1.18	0.76	1.03	YES
FANCD2	Chip 5	β-actin & GAPDH	0.76	0.81	0.72	1.07	0.95	YES
PARP1	Chip 5	β-actin & GAPDH	0.7	0.79	0.79	1.13	1.13	YES
RAD51	Chip 5	β-actin & GAPDH	0.96	0.95	0.89	0.99	0.93	YES
RAD51C	Chip 5	β-actin & GAPDH	0.73	0.94	1.02	1.29	1.40	YES
RAD51D	Chip 5	β-actin & GAPDH	0.42	0.79	0.49	1.88	1.17	NO
BARD1	Chip 5	β-actin & GAPDH	1.17	2.4	0.81	2.05	0.69	YES
BRCA1	Chip 4	β-actin & GAPDH	1.01	0.68	0.89	0.67	0.88	YES
BRCA2	Chip 4	β-actin & GAPDH	0.89	0.83	0.9	0.93	1.01	YES
MERIT40	Chip 2	β-actin	1.04	0.44	0.66	0.42	0.63	YES
BRE	Chip 4	β-actin & GAPDH	1.06	0.81	1	0.76	0.94	YES
CCDC98	Chip 4	β-actin & GAPDH	1.1	0.59	0.87	0.54	0.79	YES
RAP80	Chip 5	β-actin & GAPDH	0.6	0.75	0.6	1.25	1.00	YES
BRCC3	Chip 1	β-actin & GAPDH	1.25	0.88	1.21	0.70	0.97	YES
ERCC1	Chip 5	β-actin & GAPDH	0.55	1.17	0.66	2.13	1.20	NO
ERCC2	Chip 5	β-actin & GAPDH	1.01	0.75	1.4	0.74	1.39	YES
XRCC5	Chip 5	β-actin & GAPDH	0.93	1.05	1.04	1.13	1.12	YES
XRCC6	Chip 5	β-actin & GAPDH	0.63	0.73	0.63	1.16	1.00	YES
POLB	Chip 5	β-actin & GAPDH	0.69	1.28	0.56	1.86	0.81	YES
MLH1	Chip 5	β-actin & GAPDH	N/A	N/A	N/A	N/A	N/A	N/A
MSH2	Chip 5	β-actin & GAPDH	0.56	0.83	0.68	1.48	1.21	YES
PMS1	Chip 5	β-actin & GAPDH	0.77	0.78	0.85	1.01	1.10	YES
PIK3CA	Chip 5	β-actin & GAPDH	0.47	0.61	0.5	1.30	1.06	NO
KRAS	Chip 5	β-actin & GAPDH	0.94	0.82	1.05	0.87	1.12	YES
BRAF	Chip 4	β-actin & GAPDH	1.06	0.8	0.83	0.75	0.78	YES
CDK12	Chip 6	β-actin & GAPDH	1.22	1.03	0.9	0.84	0.74	YES
TP53	Chip 5	β-actin & GAPDH	0.64	0.93	0.75	1.45	1.17	YES
ABHD8	Chip 2	β-actin	0.46	1.04	0.59	2.26	1.28	NO
DDA1	Chip 2	β-actin	0.53	1.08	0.79	2.04	1.49	NO
GTPBP3	Chip 2	β-actin	0.59	1.21	0.79	2.05	1.34	NO
MRPL34	Chip 3	β-actin & GAPDH	0.36	0.91	0.73	2.53	2.03	NO
MYO9B	Chip 3	β-actin & GAPDH	0.9	0.82	0.76	0.91	0.84	YES
NR2F6	Chip 3	β-actin & GAPDH	0.55	1.14	0.79	2.07	1.44	NO
USE1	Chip 4	β-actin & GAPDH	0.88	0.69	0.67	0.78	0.76	YES

Table 6.12**:** Summary of the fold changes for the selected genes in the A2780CP and silenced *MERIT40* counterpart cell lines. Highlighted in green are 2 fold changes or more. Higher than 2 fold changes indicated over-expression of the gene and lower than 0.5 fold changed indicate under-expression of the genes. The column on the far right namely fold change concordance is indicating whether the fold changes observed are concordant or not.

6.4 Discussion

Platinum based chemotherapy is the first line treatment for ovarian cancer patients. Although initial treatment is effective especially in patients that carry *BRCA1* mutations most of the patients develop chemoresistance to the drugs and relapse. There are two types of chemoresistance; the intrinsic and the acquired chemoresistance. In ovarian cancer the chemoresistance is likely due to acquired chemoresistance. This mechanism of acquired chemoresistance can include the genetic or epigenetic alterations of oncogenes or tumour suppressor genes. Since MERIT40 is a BRCA1 associated protein involved in DNA repair, I hypothesised that the gain in function role I proposed in EOC development for *MERIT40* may be exerted by the encoded protein of the gene providing the cancer cells with tolerance against DNA damage and increased resistance to platinum drugs. I also investigated whether *MERIT40* would have another biological relevance to the development of EOC as an oncogene by regulating cell cycle progression, proliferation or migration.

6.4.1 Investigation of MERIT40 function in chemoresistance and DNA repair in ovarian cancer

Treatment of the EOC control and *MERIT40* knocked down cell lines with cisplatin and carboplatin did not reveal any significant differences in drug sensitivity according to the IC50s generated from the dose response curves. A slight decrease of the IC50 was observed for EFO27_M1 compared with the controls when dosed with cisplatin. It could be argued that a main limitation of this experiment was the end point MTT readout to evaluate cell viability due to cisplatin cytotoxicity might not be sensitive enough to pick up changes in cell viability when monitoring effects on DNA damage. Previous research evaluating the sensitivity of cells to DNA damage caused by irradiation when *MERIT40* is knocked down uses clonogenic assays in order to assay the effects on cell viability (Wang et al, 2009). Because of the high toxicity of cisplatin and the extreme precautions required when working with it and the numerous samples and replicates included in my study, performing clonogenic assays for this study would have been extremely difficult.

The effectiveness of cisplatin in inhibiting cell viability after *MERIT40* knockdown was assayed by calculating the IC50 based on the dose response curves generated after dosing with increasing concentrations of cisplatin. The IC50 represents the amount of cisplatin needed to kill 50% of the cells. Though comparisons of IC50s is a conventional method to assay for the effect of an inhibitor in cell growth and viability, closer step-wise inspections of the curves could reveal mechanistic information about the pathway studied providing different insights than the IC50 provides. There are no reports to my knowledge of such observations by inspecting part of dose response curves but the majority of the other studies are using 4-7 doses of platinum drugs (Fajac et al, 1996, Sevin et al, 1997, Perego et al, 1998, Asselin et al, 2001, Katano et al, 2002, Li et al, 2009, Wei et al, 2009) whereas in the present study I have performed 20 doses of cisplatin making possible the closer examination of smaller parts of the dose response curve as the doses slowly increase.

I have hypothesised that some populations of cells within the same EOC cell line could respond differently to cisplatin based on an initial observation of some patterns in dose response curves of the different EOC cell lines that were used to test cisplatin chemoresistance initially in order to select the cell lines to knock down *MERIT40*. It was apparent that dose response curves consistently showed ‘humps’ particularly in the small cisplatin doses for cell lines C13, EFO27, OVCAR3 and SCOV3IP. These patterns were consistent and reproducible between experiments and observed at exactly the same positions corresponding to specific cisplatin concentrations. These patterns were not correlated with any of the experimental conditions, such as position on the plate reader. These ‘humps’ could be indicative of a cell population within the cell lines that may have a different molecular mechanism for response to DNA damage caused by cisplatin.

Additionally, for cell lines SCOV3IP and EFO27 it is interesting that even though judging from their IC50 they appear to be quite sensitive to cisplatin, at very high cisplatin concentrations a population of cells remains resistant with the curve reaching a plateau. This may suggest that heterogeneity in EOC is responsible for the development of chemoresistance in patients and may suggest that intratumour heterogeneity has a role in acquired chemoresistance.

Cell lines SCOV3IP and EFO27 were closely examined under the microscope and the heterogeneity observed was compared to the rest of the EOC cell lines tested. SCOV3IP and EFO27 had at least 5-6 morphologically distinct clones observed (Data not shown). It will be interesting in the future to examine subclonal resistance for the different subtypes of EOC cell lines by sparsely seeding the cell lines and picking morphologically distinct colonies to test their resistance to chemotherapeutic drugs such as cisplatin.

Whether the observed differential response to cisplatin within EOC cell lines has to do with cancer stem cells (CSC) which have been suggested to have a role in intrinsic chemoresistance in several cancers (Liu et al, 2012) or different molecular switches activating different DDR pathways or a combination of both, should be further investigated. A recent study has shown that side population cells, which have the characteristics of cancer stem cells, in OVCAR3 showed increased resistance to cisplatin and other chemotherapeutic drugs exposed (Luo et al, 2012).

Inspecting the dose response curves for small, medium and high cisplatin concentrations for the EFO27 control and *MERIT40* knockdown cell lines I observed that up to 10μM of cisplatin dosing, MERIT40 deficient cell line EFO27_M1 was more sensitive to cisplatin compared to the controls (Figure 6.24). In higher dosing of cisplatin the sensitivity to cisplatin between control and knockdown cell lines was not different anymore and the dose response curves coincided. The same was not observed with the EFO27_M3 but it is not surprising as greater knockdown levels of *MERIT40* might be required for that response to take place. Similarly, A2780CP_M1 cell line seems to be more sensitive to cisplatin between 0 and 7.5 μM of cisplatin doses than the controls but in a more subtle manner than the one observed in EFO27 and should be treated with caution as the error bars were too large (Appendix 4, Figure 4). MPSCI and OVCAR3 did not show any sensitivity to cisplatin at any point of the dose response curves.

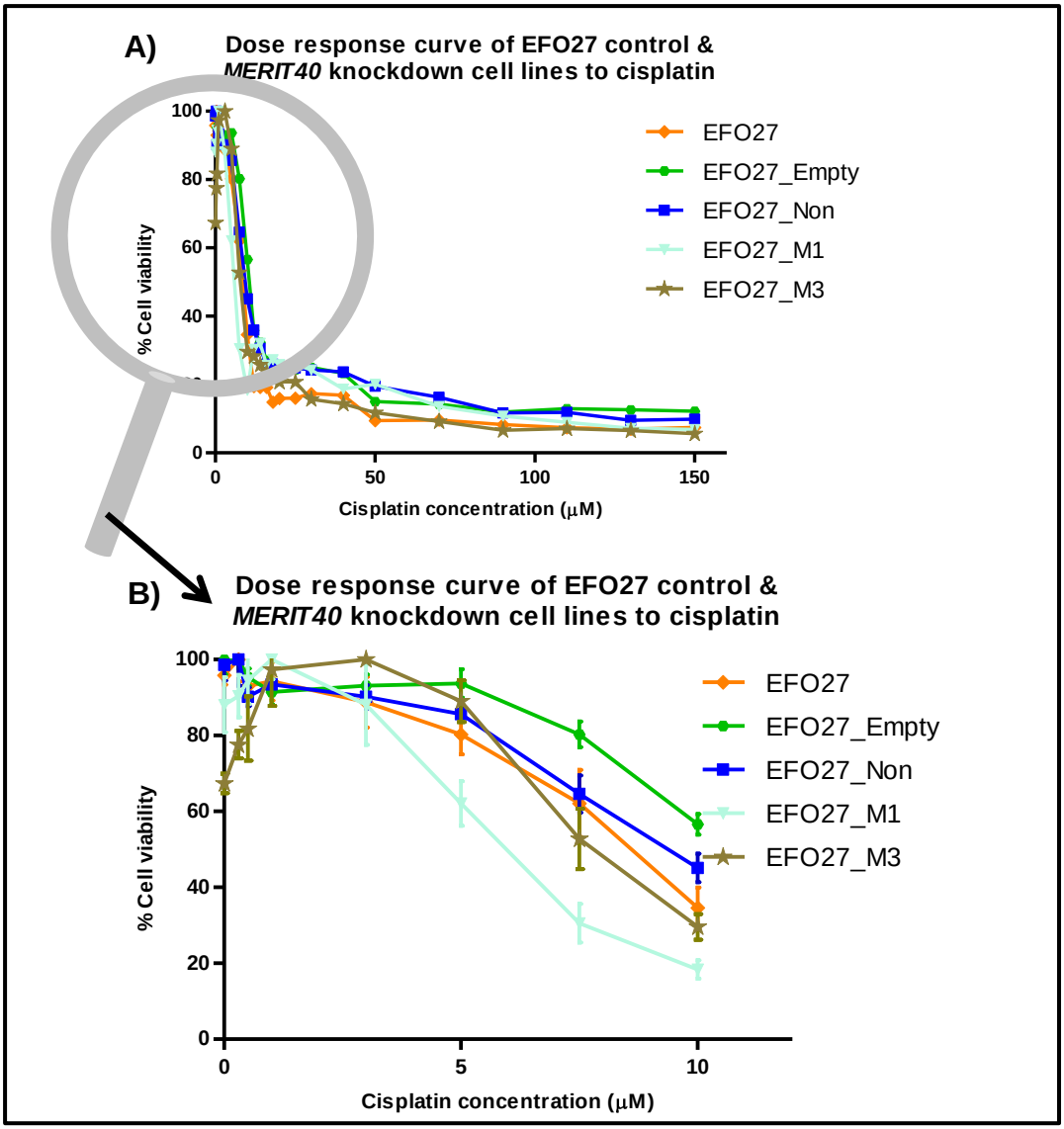


Figure 6.24**:** Close inspection of cisplatin dose response curves for EFO27 control and *MERIT40* knockdown cell lines. A) Dose response curve of EFO27 control and *MERIT40* knockdown cell lines for increasing doses of cisplatin up to 150μM did not show any differences in the cisplatin sensitivity. B) Close inspection of the initial part of the curve clearly shows the difference of the EFO27_M1 cells response up to 10μM of cisplatin dosing. The error bars are showing the standard error of mean (SEM) representing how accurate the estimate of the mean cell viability is for each cisplatin concentration.

The observation that different cisplatin concentrations seem to trigger a different response from one EOC cell line in the *MERIT40* knockdown model resemble the 'humps' of different cisplatin response observed before in EOC cell lines. Maybe they are caused by different cell populations within the cell lines responding in a different way to MERIT40 depletion. More importantly, another possible explanation would be that an alternative DNA pathway could be employed when the cells are 'threatened' with higher doses of cisplatin. The cancer cell is constantly exploiting various mechanisms in order to overcome stress and adapt to threats such as chemotherapeutic drugs. The possibility of the cell employing another DDR response to deal with damage caused by platinum drugs is very possible. PARP is participating in Base Excision Repair (BER) pathway but it has been previously shown that cells that have defects in HR they rely on PARP and BER to maintain genome integrity and repair the DNA DSBs caused by exogenous stress (reviewed by Wang and Weaver, 2012). When the BER pathway is blocked by PARP1 inhibitors in *BRCA1* and *BRCA2* mutant cells, which have defected HR, it results in cell death and this demonstrates the cooperation of the two different pathways to repair damage (Farmer et al, 2005). Based on the principle of compensating alternative mechanisms being employed as a response to platinum triggered DNA damage, it would be intriguing to find out whether dosing MERIT40 deficient cells with a combination of cisplatin and PARP1 inhibitors would cause increased cell death compared to treating the controls EOC cell lines with PARP1 inhibitors alone.

The gene expression assay I performed did not identify any differential expression of *PARP1* between the cell lines when MERIT40 was depleted or not. However I found *PARP1* to be significantly over-expressed in EOC cell lines compared to NOSE & FTE cell lines ($P=1.6\times10^{-9}$), (Appendix 4, Figure 5). It would be interesting to assay *PARP1* expression in the EOC control and EOC *MERIT40* knockdown cell lines after they are treated with doses of cisplatin corresponding to before and after the dose response curves start to coincide. If PARP1 is employed an as alternative DNA repair pathway after certain doses of cisplatin then it may be more highly expressed at the doses where the dose response curves were found to coincide. It has been shown that in the event of

DNA SSBs and DSBs the catalytic activity of PARP1 is increased by 100fold (Bouchard et al, 2003) but to my knowledge there are no reports on whether the transcript output for *PARP1* increases too.

Another question raised by this study is why these EOC cell lines respond differently to cisplatin when *MERIT40* was knocked down and one answer could be because they represent different EOC subtypes. It is interesting that high and low grade serous cell lines, MPSC1 and OVCAR3 respectively, did not show any increased sensitivity to cisplatin when *MERIT40* was knocked down but the mucinous cell line EFO27 did show that initial improved response to low doses of cisplatin and also a small decrease in the IC50 when *MERIT40* is knocked down compared to the controls however with very large 95% confidence intervals. A small initial response up to 7.5μM was also shown by A2780CP which is an undifferentiated tumour cell line. It is curious that the EOC risk association found by the GWAS in chromosome 19 with a coding SNP within *MERIT40* was driven by serous subtypes only (Bolton et al, 2010). Therefore it would be expected that any functional role of *MERIT40* would have been more apparent in serous EOC cell lines. However, as previously discussed, a SNP in *MERIT40* associated with susceptibility for serous EOC does not necessarily indicate that the gene is not somatically involved in EOC development, progression and survival of the disease in other subtypes. I investigated the publicly and in house available information for these four EOC cell lines to see if I could find any differences. I have evaluated the expression of *BRCA1* between the four EOC cell lines (data not shown) and I found no differences. *PARP1* expression between them also did not differ.

p53 is known to be involved in drug induced apoptosis and has been previously shown to suppress HR by transcriptionally regulating RAD51 (Lopez et al, 2006). Thus cells with depleted p53 could be more resistant in cisplatin induced cytotoxicity avoiding apoptosis of unrepaired damage caused by the drugs. Additionally, mutations in *TP53* have been shown to have a role in cisplatin resistance in ovarian cancer (Perego et al, 1996). This raises questions regarding the role of p53 in the response of cancer cell lines to DNA crosslinking cytotoxic agents. Recent work has shown that whereas cells that are p53 deficient have suppressed HR activity, when treated with an intrastrand

crosslinking agent mitomycin C, their deficiency in p53 did not impair the cells ability to use HR for the removal of DBS (Sirbu et al, 2011). An older study in ovarian cancer, however, has shown that cells deficient in p53 demonstrate resistance to cisplatin but when treated with a novel platinum agent DACH-aceto-Pt the p53 protein production was induced and drug related apoptosis was triggered (Hagopian et al, 1999). Mutational analyses performed for those cell lines report *TP53*, *PTEN* and *MSH2* mutations in EFO27. *TP53* mutations were present also in the OVCAR3 cell line (www.sanger.ac.uk). A2780CP has been shown to harbour *TP53* mutations (Song et al, 1997) which result in the expression of mutant conformations of p53 that are transcriptionally inactive. Based on that information it could be suggested that the p53 status of the cell lines EFO27, OVCAR3 and A2780CP did not seem to be the differentiating factor for the subtle differences they showed in cisplatin sensitivity when *MERIT40* was knocked down.

Collectively, the known or observed information I have on the 4 EOC cell lines used in this study, do not provide any compelling discriminating characteristics on why they should be more sensitive to cisplatin when *MERIT40* is knocked down. The only difference is the mutation in *MSH2* found only in EFO27 that could be treated as a discriminating factor. Therefore, I investigated this further. In the expression assay I performed I found that whereas *MSH2* was significantly over-expressed in EOC compared to NOSE cell lines ($P=2.5\times10^{-9}$), EFO27 had at least two fold less expression of *MSH2* compared to the other 3 cell lines that were used to knock down *MERIT40* (Appendix 4, Figure 6 A & B respectively). However, this finding does not correlate with known information regarding the role of MSH2 in chemoresistance for it to be deemed responsible for the observed difference in response to low doses of cisplatin between the 4 cell lines. Platinum complexes have been shown in some cancer types to interfere with normal activity of the mismatch repair (MMR) and prevent the DDR leading to apoptosis. The MMR has been previously shown to be implicated with repair of DNA damage caused by cisplatin as the loss of MMR activity conferred resistance to cisplatin to endometrial and colon cancer cells (Aebi et al, 1996, Fink et al, 1998). In ovarian cancer the role of MMR protein in cisplatin resistance is more

controversial. In a study of 75 ovarian carcinomas for microsatellite instability and mRNA expression of members of the MMR pathway, loss of *MSH2* and other members of the pathway, was not correlated with cisplatin resistance (Helleman et al, 2006). Additionally, the immunohistochemical expression of MLH1 and MSH2 proteins was not different between platinum resistant and platinum sensitive ovarian cancer patients (Magnowska et al, 2008). Other studies have reported however that ovarian cancer cells that overexpress MMR proteins are sensitive to cisplatin (Martin et al, 2008).

It should be noted that observed differences in results from different studies using EOC cell lines may have resulted from mutations and other alterations that could have occurred during serial propagation of cells in culture leading to cell lines no longer representing reliable models of their original source material. Such alterations have been previously shown for EOC cell lines (Hughes et al, 2007). This is a limitation that should be mentioned for all the studies that used EOC cell lines including this study.

Based on all these data, unfortunately, in this study I have not managed to fully unravel the role of MERIT40 in DNA repair triggered by cisplatin, but the findings suggest that alternative pathways may be activated by the cell when a DDR pathway is depleted and that a sub-population of cells may be conferring increased resistance in EOC.

I have also investigated the role of MERIT40 in DNA damage repair caused by ionising irradiation (IR) and spontaneous damage caused during replication. To my knowledge this is the first study that reports that *MERIT40* knockdown in mucinous ovarian cancer cells contributes to increased cell death after ionizing radiation although the result is moderate and should be treated with caution. More importantly, I observed that unrepaired DNA from spontaneous damage during replication was significantly increased as indicated by increased levels of phosphorylated γH2AX nuclear foci when *MERIT40* was knocked down in the EFO27 mucinous EOC cell line. It is likely that MERIT40 may be more importantly involved in HR employed by the cancer cell for surviving replication errors rather than for surviving aggressive exogenous DNA damage induced by platinum drugs, in which case alternative repair pathways may be used.

6.4.2 Investigation of *MERIT40* gain of function role in EOC development

Based on the indication that MERIT40 may contribute in EOC development as part of the HR response to replication induced DNA DSBs rather than chemoresistance I wanted to see if MERIT40 is involved through another function in oncogenesis. Previous work has shown that only a fraction of cellular BRE/MERIT40 and BRCC36 localizes to the sites of DNA breaks, raising the possibility that BRE/ MERIT40 and BRCC36 may also have other functions in the cell (Feng et al, 2009). I have therefore chosen to investigate whether silencing of *MERIT40* in EOC cell lines would cause the reversal of neoplastic phenotypes. I investigated the effect of *MERIT40* knockdown on cell cycle regulation, proliferation, migration and anchorage independent growth.

The most significant finding was the decrease of cells in G2/M phase observed in mucinous cell line EFO27 with depleted *MERIT40*. An interesting observation is the 50% decrease observed in cell lines EFO27_M1 and EFO27_M3 for the cells with 8N content. A reduction in the cells with 8N DNA content may suggest that the arrest observed was not due to reduced proliferation but due to reduction in ploidy. In fact EFO27 has been shown to be hypotretraploid (Lambros et al, 2005). In a follow up analysing metaphase spreads I have found that the amount of polyploid cells decreases when *MERIT40* is knocked down but the amount of cells available to analyse was not sufficient to suggest significance for the observed difference. Recent work has demonstrated that polyploid cells need intact DDR pathways to overcome replication stress induced barriers for tumour progression. This showed that polyploidy in cells is causing the over-expression of DNA repair genes including *BRCA1* (Zheng et al, 2012). Thus it is possible that the reduction in ploidy and/or arrest of EFO27 cells with depleted MERIT40 in cell cycle checkpoints are a result of defective DNA repair caused by the absence of MERIT40.

Another characteristic of transformed cells is their ability to grow in an anchorage independent manner. I have therefore investigated whether MERIT40 has a role in ovarian cancer development by promoting malignant transformation of cells after increasing their capability to proliferate independently of external and internal signals losing the normal contact inhibition response. Knocking down *MERIT40* did not cause a reduction in the

anchorage independent growth ability for cell lines EFO27, MPSC1 and A2780CP. I have observed a small reduction in the number of colonies formed when *MERIT40* was knocked down in the high grade serous EOC cell line OVCAR3. Previous work in breast cancer has identified a negative regulator of BRCA1 that when it is over-expressed is associated with anchorage independent growth thus associating the over-expression of *BRCA1* with increased ability to proliferate anchorage independently (Welcsh et al, 2001). However, I found that the expression for *BRCA1* in OVCAR3 control and *MERIT40* knockdown cell lines was not different (data not shown) so it is more possible that MERIT40 could promote anchorage independent growth by regulating another pathway involved in the generation of that particular neoplastic phenotype. According to the TCGA database, amplifications in the *MERIT40* locus, that may correlate to higher expression of the gene have a significant effect on the levels of several proteins with the most significant being a large reduction in the protein levels of MAPK9 (alternatively named *JNK2*) (P= 0.0008). That may form an interesting correlation with the decreased growth in soft agar observed in OVCAR3 EOC cell lines with depleted MERIT40. In a study investigating non-small cell lung carcinoma tumorigenesis it was shown that cells deficient in the α isoform of MAPK9 had decreased anchorage independent growth (Nitta et al, 2011). It is possible that MERIT40 acts as a regulator of the activity of MAPK9, thus dysregulation of MERIT40 expression levels may cause increased activity of MAPK9 leading to the cells becoming less dependent on ECM adhesion and able to grow in an anchorage independent manner leading to tumorigenesis in serous EOC.

6.4.3 Evaluating the effects of *MERIT40* knockdown on regulating other pathways involved in tumorigenesis

In an attempt to identify additional pathways that MERIT40 may be regulating, I have chosen to investigate whether knocking down *MERIT40* would affect the mRNA expression levels of several genes that are involved in cell cycle regulation, chromosome segregation, proliferation, differentiation and DNA repair pathways.

For EFO27 cell line none of the selected genes was found to be differentially expressed when MERIT40 was depleted thus I propose that the phenotypic effect that MERIT40 may have in cell cycle regulation is exerted by its own role in DNA repair and not via regulating another pathway, at least not by being involved in the transcriptional regulation of components involved in malignant transformation pathways.

For MPSC1 I found that *BRAF* expression was significantly reduced when MERIT40 was depleted. That was a very interesting observation because the low grade serous EOC cell line MPSC1 is already reported to have activating *BRAF* mutations (www.sanger.ac.uk). *BRAF* mutations have been implicated with early EOC development and these mutations are associated with the MAPK pathway in ovarian low grade serous tumours and the constitutive activation of the MAPK pathway plays a key role in the development of low grade serous tumours containing mutant *KRAS* or *BRAF* (Pohl et al, 2005). It is conceivable that MERIT40 somehow attenuates the increased expression of BRAF in low grade serous tumours with activating mutations of *BRAF* and contributing in such way to activation of the MAPK pathway leading to increased proliferation and decreased apoptosis. The decreased expression of *BRAF* in MPSC1 with depleted MERIT40 did not relate to any of the phenotypic assays performed to assess for reversal of the neoplastic phenotype. I have not found that knocking down *MERIT40* in MPSC1 had any effects reducing the proliferation after evaluating the cells in the S phase of the cell cycle based on BrDU incorporation for the control and knockdown cell lines. However, a decrease in the proliferation was assumed in the initial evaluation of cell proliferation with the MTT assay. The fact that this was not validated by the more specific BrDU analysis may be explained by the possibility that the observation with MTT was not due to an effect of MERIT40 in the proliferative ability for the cells but in their ability to undergo apoptosis. Additional work researching the effect of *MERIT40* knockdown in apoptosis could be beneficial into understanding the possible role of MERIT40 in EOC development by attenuating the expression of BRAF leading to activation of the MARK pathway.

For OVCAR3 I found that when MERIT40 was depleted *GTPBP3* expression was significantly increased and *ABHD8* was significantly reduced

compared to the control cell line. Both of those genes are located in locus 19p13 within 100kb downstream of *MERIT40* and to my knowledge they have not been implicated with any functions in carcinogenesis. Their function is not clearly understood yet but is believed that *ABHD8* may code for a protein that has catalytic and hydrolase activity involved in metabolic processes and GTPBP3 for a protein that may play a role in mitochondrial tRNA modification (www.genecards.org).

The significant reduction in *RAD51D* expression when *MERIT40* was knocked down in the OVCAR3 cell line was an interesting observation. RAD51D is involved in HR as part of the RAD51 complex which searches the genome for an intact copy of the damaged DNA on the sister chromatid for HR to take place. MERIT40 is implicated in HR by stabilising the BRCA1-Rap80-Abraxas complex in DNA DSBs and has mainly been shown to be involved in the recognition and initiation of the process as previously described. My observation could indicate an additional role of MERIT40 in HR by positively regulating the crucial scanning function of RAD51. Recently *RAD51* has been proposed to be a gene that confers moderate risk for the development of EOC. Inactivating mutations in *RAD51D* have been shown to confer susceptibility to ovarian cancer and mutations of this gene have also been found in ovarian cancer families with no *BRCA1* or *BRCA2* mutations (Loveday et al, 2011, Osher et al, 2012). Additionally, cells that are deficient in RAD41D have been shown to be sensitive to treatment with PARP inhibitors (Loveday et al, 2011). Hypothesising that the activity of RAD51D may be regulated by MERIT40 levels it would be interesting to investigate whether OVCAR3 with depleted MERIT40 would be more sensitive to PARP inhibitors compared to the control cell line.

Finally, I found that A2780CP with *MERIT40* knocked down, had increased expression of *BARD1*. BARD1 is involved in DNA repair and HR by forming a complex with BRCA1 which is involved in the activation of the G1/S, S-phase and G2/M checkpoints but is a process that is not yet well characterised. When the complex BRCA1-BARD1 is depleted then the activation of the G1/S checkpoints in response to damage by ionising radiation is compromised (Fabbro et al, 2004). Additionally, the G2/M checkpoint was also found defective in cells lacking BARD1 (reviewed in Roy et al, 2012). My

observation could be indicative of MERIT40 being involved in checkpoint control, possibly the G2/M checkpoint, by negatively regulating BARD1 causing the cells to avoid arresting in checkpoint controls where they would undergo apoptosis but continue to proliferate uncontrollably leading to neoplastic transformation in EOC. However, this hypothesis was not reflected by any of the phenotypic assays performed. One limitation of this study is that the expression of genes at the mRNA level does not guarantee that an active protein is produced such is the case previously mentioned with A2780CP expressing *TP53* on the mRNA level but the protein expressed is inactive. Thus the results should be treated with caution and can only be used to speculate possible interactions of MERIT40 with other neoplastic pathways.

6.4.4 Conclusion

I observed that in low doses of cisplatin EFO27 mucinous EOC cell line became more sensitive when MERIT40 was depleted. That could be due to either a subpopulation of cells that respond differently to cisplatin and/or alternative repair pathways being employed by the cancer cell to ensure that it would regain resistance towards higher doses of cisplatin. Additionally, I propose that MERIT40 may be involved in HR employed for the DNA repair of DSBs induced during replication and that might show a potential role for the gene in EOC early stages of development rather than chemoresistance. Consistent with this hypothesis is the observation that knocking down *MERIT40* caused cell cycle arrest and reduction of ploidy for the EFO27 cell line which may be just another response to unrepaired DNA damage caused by replicative stress. It could be interesting in the future to use NOSE and FTE 3D *in vitro* models and overexpress *MERIT40* in the absence and presence of p53 in order to study the potential effect of this gene in the early stages of EOC development.

Undoubtedly, the DDR pathways that contribute towards the cytotoxicity to cisplatin and the mechanisms which are employed by the cancer cell are several. These alternative mechanisms can modulate the cellular response to cisplatin induced DNA damage in a tightly coordinated manner and may also vary between the different subtypes of EOC. The possibility of *MERIT40*

knockdown contributing to deficiency in HR leading to increased EOC cell death after cisplatin exposure remains a point of debate and should be further investigated in relation to alternative DDR pathways. Further research investigating the mechanisms that are responsible for the collaboration of all the recognised DDR pathways in cisplatin resistance for the different subtypes of EOC would be valuable and could lead and even more complex and individualised combination therapies to overcome resistance to chemotherapy.

Finally, it still remains under debate whether MERIT40 has other functions in the cell than being involved with DNA repair. Over-expression of *MERIT40* being apparent in EOC cell lines suggests that MERIT40 may contribute to EOC development through other pathways leading to neoplastic transformation. However, I did not find any really compelling information supporting this hypothesis. Additional work investigating other oncogenic pathways that MERIT40 may be involved for the development of different EOC subtypes would be critical to fully understand the function of that gene in EOC development.

7 Conclusions and Future work

Epithelial ovarian cancer (EOC) is the most common cause of gynaecological cancer deaths but is still poorly understood. The origin of ovarian cancer has been considered to be the normal Ovarian Surface Epithelium (NOSE), but a proportion of ovarian cancers may also originate from Fallopian Tube Epithelial (FTE) cells. Additionally, the known high risk susceptibility genes for EOC do not explain all the heritable components of EOC's. Pathway based association studies and genome wide association studies aim to identify alleles of low/moderate risk that may account for the development of the remaining familial cases and sporadic EOC.

The aim of this study was to investigate the functional role of low/moderate susceptibility SNPs and candidate genes that emerge from candidate gene approaches and GWAS using appropriate tissue and cell line models. Part of this study focused in establishing and characterising a NOSE primary cell bank (n= 57), and 5 primary FTE cell lines to compare their expression with EOC cell lines in post-GWAS functional characterisation studies. Additionally, I have established 3D FTE cultures for *in vitro* models to study EOC transformation. I found that FTE3D cultures are a better biological *in vitro* model than 2D as they better reflect the *in vivo* architecture than 2D cultures. Both FTE and NOSE 3D models could be used to impose defined genetic modifications that are characteristic for different subtypes of EOC in order to elucidate the molecular events that take place for EOC initiation and development to occur.

A candidate gene approach identified nine genes (*AIFM2*, *AKTIP*, *AXIN2*, *CASP5*, *FILIP1L*, *RBBP8*, *RGC32*, *RUVBL1* and *STAG3)* using an *in vitro* model of microcell mediated chromosome 18 transfer. I tested 301 invasive EOC tumours and found frequent LOH for tagging SNPs in those genes with the highest frequencies identified for *AXIN2, AKTIP* and *RGC32* and one SNP (rs1637001) in *STAG3* showed significant preferential loss of the common allele. Also, patients with LOH in genes *AXIN2, CASP5, RRBBP8* and *AIFM2* had significantly worse survival than cases without LOH but the result was associated with stage in *AXIN2, CASP5* and *AIFM2*.

Six loci - 2q31, 3q25, 17q21, 8q24, 9p22 and 19p13 - containing susceptibility SNPs were identified in an ovarian cancer GWAS and I compared the expression of candidate genes within the risk associated loci between normal (NOSE & FTE) and EOC cell lines and found compelling evidence for the somatic role of some of these genes in EOC development. These genes included *PVT1*, *SP2*, *CBX1*, *PNPO*, *HAUS8*, *USE1*, *SKAP1*, *MERIT40* and members of the *HOXB* family of genes with an observed gain of function role and *TIPARP* and *BNC2* with an observed loss of function role in EOC development. Additionally I have investigated whether the associated SNPs would have a functional regulatory role in EOC and found weak associations for the rs9303542 genotype with expression of genes *CBX1, SNX11 and SP2 and* for the rs2072590 genotype with *HOXD1 expression.* Additionally, I found compelling evidence for the rs9303542 genotype being associated with *HOXB5* methylation in a cohort of 256 healthy controls. These results indicate that those genes may be the susceptibility targets of these loci regulated by the identified variants. Collectively these results also suggest the use of the NOSE & FTE versus EOC cell lines expression model as a reliable first line tool for evaluation of candidate genes' role in EOC.

I further knocked down *MERIT40* in 4 EOC cell lines and studied their chemosensitivity towards cisplatin and carboplatin, but did not find compelling evidence of a role or MERIT40 in drug resistance. However, I have found that MERIT40 depletion leads to increased accumulation of spontaneous DNA damage and cell cycle arrest characterised by reduction in ploidy of the mucinous EOC cell line EFO27

Most association studies report associated SNPs with disease but no evidence to suggest which candidate gene in a region may be functionally important as a target. This study has provided substantial evidence that GWAS are an incredibly powerful tool for understanding the relationship between carcinogenesis and genetic markers. Additionally, that a methodical approach conducting post-GWAS functional characterization of the cancer risk loci is now essential to provide us with valuable biological insights that can only be beneficial for creating more effective strategies for the screening, prevention and treatment of the disease. The post-GWAS functional characterisation

performed in this study has only looked at a small number of possible functions for the candidate genes and variants in the associated loci and much more can be investigated.

Several questions remain unanswered and require large collaborative efforts to form a conclusive picture. In the first instance, ongoing work is being carried out by members of the OCAC, including my group, to fine map the candidate loci. This involves large scale genotyping of samples using the 1000 Genomes project data in order to identify other SNPs in LD with the original SNPs in the regions of the EOC associated loci that may confer higher risk of EOC susceptibility. A well-structured approach to study the function of the candidate SNPs and fine mapping SNPs in more detail could be performed based on the findings from this thesis and additional research in the following way:

In case a SNP is a non synonymous SNP changing the amino acid sequence of a protein, further bioinformatics analysis can be performed to identify whether the changed amino acid influences protein-protein interactions, DNA binding, and if the SNP is within or likely to affect activation domains. Bionformatic analysis can be performed to determine whether coding and non coding variants are located in known promoter regions, or within splicing regions, or in regions exhibiting copy number variations. Furthermore, the SNPs could affect miRNA binding sites.

The next step is to evaluate the potential *cis*-regulatory effects of the identified variants by investigating whether the SNP affects transcriptional output or methylation of genes within close proximity (1Mb from start or distal end). In this study I have established that a reliable assay to investigate function of the candidate SNPs is looking at gene expression differences relative to NSP genotype, using the cells of origin for EOC (NOSE&FTE) since gene expression is tissue specific. However, although using such a model is the most appropriate way for reliable results its limitation was in the number of the samples. Additional work is being performed by my supervisors, Dr Susan Ramus and Prof Simon Gayther at USC, and collaborators using SNPs from the original GWAS and fine mapping SNPs, to establish whether there are gene expression differences between the genotypes of SNPs after genotyping the

NOSE & FT cells established in this project as well as DNA from lymphoblastoid cell lines to increase the power of the study to detect significant associations. The expression of miRNAs in the regions of susceptibility should also be examined relative to the genotype of candidate SNPs as the variants may regulate gene expression via miRNA regulation. Allele specific LOH could also be investigated for candidate SNPs that were found associated with gene expression and could provide insights to the mechanisms by which a risk allele may affect the transcriptional output of an oncogene or tumour suppressor gene based on the differential expression observed in normal versus EOC cell lines. Finally, genotype specific expression of alternative splice variants of candidate genes could be evaluated to test whether these SNPs are regulating the genes by affecting their splicing patterns.

Furthermore, the *trans* effects of the candidate SNPs could be investigated in the future by performing microarray analysis for global gene expression in NOSE and FTE cell lines as well as evaluating the methylation of genes using genome wide methylation data relative to the genotypes of the candidate SNPs.

In post GWAS-characterisation the genes that are regulated by the genotype of candidate SNPs may be the target susceptibility genes. However genes may have a somatic role in EOC development. This study has demonstrated that by evaluating the differential expression of candidate genes within the identified loci between normal (NOSE &FTE) and EOC cell lines it is possible to suggest a loss or gain of function role for these genes in EOC development. This could be validated in the future by investigating the protein expression for those genes by performing immunohistochemistry in tissue arrays and comparing between primary ovarian tumours and normal OSE and FTE tissues. Further mutational and copy number variation for this genes is now available by TCGA and could be used to correlate the findings of the expression differences observed.

It should be noted that the differential expression model used in this study mainly indicates the somatic role of the candidate genes in EOC development but probably is not indicative of the genes' function in early oncogenesis. Once the best candidate genes are established, from this study

and after the fine mapping approach, a coordinated study studying the functional role of these genes in the appropriate FTE, NOSE or EOC *in vitro* models would be essential to identify what are the molecular pathways these genes are using to lead to neoplastic transformation and also in regards to earlier or later stages of oncogenesis for the different subtypes. Knocking down potential oncogenes in different subtypes of EOC cell lines and assaying for neoplastic phenotype reversal can provide information about the candidate genes' function in the development and progression of the different EOC subtypes. Overexpressing potential oncogenes or knocking down potential tumour suppressor genes in 3D NOSE (Lawrenson et al, 2011) and the FTE 3D in vitro model established in his study and assaying for neoplastic transformation could provide information about development of the different EOC subtypes in early stages of transformation.

The necessity to use the appropriate *in vitro* models to study early or late stages of EOC development is strengthened by an interesting and puzzling observation made in this study regarding the role of tumour suppressor genes in EOC. Tumour suppressor genes such as *BRCA1*, *BRCA2* and *TP53,* all known to have a role in DNA repair and cell cycle regulation, were found to be over-expressed in EOC cells consistent with a similar trend for *BRCA1* and *BRCA2* observed in SOC based on TCGA data analysed. This could be an indication that the results obtained from the NOSE&FTE versus EOC differential expression model may describe later stages in EOC development, where the cancer cell has employed mechanisms that traditionally the cell uses to avoid neoplastic transformation, in order to overcome DNA damage from endogenous or exogenous stress such as replication errors or cytotoxic drugs respectively. Thus, tumour suppressor genes involved in such pathways may be found over expressed in cancer cells during these stages of cancer development. Based on such hypothesis the same tumour suppressor gene conferring protection to the cell before cancer initiation, is then deactivated by a somatic alteration which favours neoplastic development and later, somehow the cancer cell manages to restore its function to promote its growth and avoid DNA damage triggered apoptosis. Thus, it may be conceivable that to study the function of such a tumour suppressor genes in later stages of EOC may require knocking it down

in EOC cell lines. This hypothesis is consistent with some studies that have shown findings to support this theory, including *BRCA1* secondary activating mutations in cisplatin resistant tumours (Swisher et al, 2008) and *BRCA1* over-expression promoting cisplatin resistance in ovarian cancer cells (Chock et al, 2010). I therefore hypothesise that tumour suppressor genes may adopt "oncogenic"-like properties in cancer development by being reactivated and promoting the survival of the cancer cell may be one of the hallmarks for cancer progression. Studying the mechanisms by which such a fundamental functional switch could be achieved may provide valuable insights on cancer cells' ability to acquire chemoresistance or progress into more advances stages. Further research elucidating the role of such TSGs in later stages of cancer progression may be the key to provide novel individualised therapeutic targets and for better prognosis of late stage diagnosed EOC patients.

To conclude, I propose that the functional investigation of GWAS identified risk or survival associated loci using the models described in a step wise process that I have established in this study should lead to a better understanding of the molecular players involved in EOC initiation, development and survival and could lead to more reliable screening and individualised treatments for EOC.

8 References

AARNIO, M., SANKILA, R., PUKKALA, E., SALOVAARA, R., AALTONEN, L. A., DE LA CHAPELLE, A., PELTOMAKI, P., MECKLIN, J. P. & JARVINEN, H. J. 1999. Cancer risk in mutation carriers of DNA-mismatch-repair genes. *Int J Cancer,* 81**,** 214-8.

ABDOLLAHI, A., BAO, R. & HAMILTON, T. C. 1999. LOT1 is a growth suppressor gene down-regulated by the epidermal growth factor receptor ligands and encodes a nuclear zinc-finger protein. *Oncogene,* 18**,** 6477-87.

ACIEN, P. 1992. Embryological observations on the female genital tract. *Hum Reprod,* 7**,** 437-45.

ADZHUBEI, I. A., SCHMIDT, S., PESHKIN, L., RAMENSKY, V. E., GERASIMOVA, A., BORK, P., KONDRASHOV, A. S. & SUNYAEV, S. R. 2010. *A method and server for predicting damaging missense mutations*, Nat Methods. 2010 Apr;7(4):248-9.

AEBI, S., KURDI-HAIDAR, B., GORDON, R., CENNI, B., ZHENG, H., FINK, D., CHRISTEN, R. D., BOLAND, C. R., KOI, M., FISHEL, R. & HOWELL, S. B. 1996. Loss of DNA Mismatch Repair in Acquired Resistance to Cisplatin. *Cancer Research,* 56**,** 3087-3090.

AGARWAL, R. & KAYE, S. B. 2003. Ovarian cancer: strategies for overcoming resistance to chemotherapy. *Nat Rev Cancer,* 3**,** 502-516.

AGGELER, J., WARD, J., BLACKIE, L. M., BARCELLOS-HOFF, M. H., STREULI, C. H. & BISSELL, M. J. 1991. Cytodifferentiation of mouse mammary epithelial cells cultured on a reconstituted basement membrane reveals striking similarities to development in vivo. *J Cell Sci,* 99**,** 407-17.

AGORASTOS, T., MASOURIDOU, S., LAMBROPOULOS, A. F., CHRISAFI, S., MILIARAS, D., PANTAZIS, K., CONSTANTINIDES, T. C., KOTSIS, A. & BONTIS, I. 2004. P53 codon 72 polymorphism and correlation with ovarian and endometrial cancer in Greek women. *Eur J Cancer Prev,* 13**,** 277-80.

AHMED, A. A., ETEMADMOGHADAM, D., TEMPLE, J., LYNCH, A. G., RIAD, M., SHARMA, R., STEWART, C., FEREDAY, S., CALDAS, C., DEFAZIO, A., BOWTELL, D. & BRENTON, J. D. 2010. Driver mutations in TP53 are ubiquitous in high grade serous carcinoma of the ovary. *J Pathol,* 221**,** 49-56.

AHMED, N., MAINES-BANDIERA, S., QUINN, M. A., UNGER, W. G., DEDHAR, S. & AUERSPERG, N. 2006. Molecular pathways regulating EGF-induced epithelio-mesenchymal transition in human ovarian surface epithelium. *American Journal of Physiology - Cell Physiology,* 290**,** C1532-C1542.

AHMED, S., THOMAS, G., GHOUSSAINI, M., HEALEY, C., HUMPHREYS, M., PLATTE, R., MORRISON, J., MARANIAN, M., POOLEY, K., LUBEN, R., ECCLES, D., EVANS, D., FLETCHER, O., JOHNSON, N., DOS SANTOS SILVA, I., PETO, J., STRATTON, M., RAHMAN, N., JACOBS, K., PRENTICE, R., ANDERSON, G., RAJKOVIC, A., CURB, J., ZIEGLER, R., BERG, C., BUYS, S., MCCARTY, C., FEIGELSON, H., CALLE, E. & THUN, M. 2009. Newly discovered breast cancer susceptibility loci on 3p24 and 17q23.2. *Nat Genet,* 41**,** 585 - 590.

AIDA, H., TAKAKUWA, K., NAGATA, H., TSUNEKI, I., TAKANO, M., TSUJI, S., TAKAHASHI, T., SONODA, T., HATAE, M., TAKAHASHI, K., HASEGAWA, K., MIZUNUMA, H., TOYODA, N., KAMATA, H., TORII, Y., SAITO, N., TANAKA, K., YAKUSHIJI, M. & ARAKI, T. 1998. Clinical features of ovarian cancer in Japanese women with germ-line mutations of BRCA1. *Clin Cancer Res,* 4**,** 235-40.

AKAGI, T., ITO, T., KATO, M., JIN, Z., CHENG, Y., KAN, T., YAMAMOTO, G., OLARU, A., KAWAMATA, N., BOULT, J., SOUKIASIAN, H. J., MILLER, C. W., OGAWA, S., MELTZER, S. J. & KOEFFLER, H. P. 2009. Chromosomal abnormalities and novel disease-related regions in progression from Barrett's esophagus to esophageal adenocarcinoma. *International Journal of Cancer,* 125**,** 2349-2359.

AL-DHAHERI, M. H., SHAH, Y. M., BASRUR, V., PIND, S. & ROWAN, B. G. 2006. Identification of novel proteins induced by estradiol, 4-hydroxytamoxifen and acolbifene in T47D breast cancer cells. *Steroids,* 71**,** 966-978.

ALEXEI, J. Y. A. A. V. T. V. 2010. The roles of PARP1 in gene control and cell differentiation. *Curr Opin Genet Dev.,* 20**,** 512–518.

AMOS, C. I., WU, X., BRODERICK, P., GORLOV, I. P., GU, J., EISEN, T., DONG, Q., ZHANG, Q., GU, X., VIJAYAKRISHNAN, J., SULLIVAN, K., MATAKIDOU, A., WANG, Y., MILLS, G., DOHENY, K., TSAI, Y. Y., CHEN, W. V., SHETE, S., SPITZ, M. R. & HOULSTON,

R. S. 2008. Genome-wide association scan of tag SNPs identifies a susceptibility locus for lung cancer at 15q25.1. *Nat Genet,* 40**,** 616-22.

ANTONIOU, A., PHAROAH, P. D., NAROD, S., RISCH, H. A., EYFJORD, J. E., HOPPER, J. L., LOMAN, N., OLSSON, H., JOHANNSSON, O., BORG, A., PASINI, B., RADICE, P., MANOUKIAN, S., ECCLES, D. M., TANG, N., OLAH, E., ANTON-CULVER, H., WARNER, E., LUBINSKI, J., GRONWALD, J., GORSKI, B., TULINIUS, H., THORLACIUS, S., EEROLA, H., NEVANLINNA, H., SYRJAKOSKI, K., KALLIONIEMI, O. P., THOMPSON, D., EVANS, C., PETO, J., LALLOO, F., EVANS, D. G. & EASTON, D. F. 2003. Average risks of breast and ovarian cancer associated with BRCA1 or BRCA2 mutations detected in case Series unselected for family history: a combined analysis of 22 studies. *Am J Hum Genet,* 72**,** 1117-30.

ARIAS, A. M. 2001. Epithelial mesenchymal interactions in cancer and development. *Cell,* 105**,** 425-31.

ARNESON, N., HUGHES, S., HOULSTON, R. & DONE, S. 2010. Whole-Genome Amplification by Degenerate Oligonucleotide Primed PCR (DOP-PCR). *Cold Spring Harbor Protocols,* 2008**,** pdb.prot4919.

ARNOLD, N., HAGELE, L., WALZ, L., SCHEMPP, W., PFISTERER, J., BAUKNECHT, T. & KIECHLE, M. 1996. Overrepresentation of 3q and 8q material and loss of 18q material are recurrent findings in advanced human ovarian cancer. *Genes Chromosomes Cancer,* 16**,** 46-54.

ASSELIN, E., MILLS, G. B. & TSANG, B. K. 2001. XIAP regulates Akt activity and caspase-3-dependent cleavage during cisplatin-induced apoptosis in human ovarian epithelial cancer cells. *Cancer Res,* 61**,** 1862-8.

ATTANASIO, C., REYMOND, A., HUMBERT, R., LYLE, R., KUEHN, M., NEPH, S., SABO, P., GOLDY, J., WEAVER, M., LEE, K., HAYDOCK, A., DERMITZAKIS, E., DORSCHNER, M., ANTONARAKIS, S. & STAMATOYANNOPOULOS, J. 2008. Assaying the regulatory potential of mammalian conserved non-coding sequences in human cells. *Genome Biol,* 9**,** R168.

AUERSPERG, N., EDELSON, M. I., MOK, S. C., JOHNSON, S. W. & HAMILTON, T. C. 1998. The biology of ovarian cancer. *Semin Oncol,* 25**,** 281-304.

AUERSPERG, N., PAN, J., GROVE, B. D., PETERSON, T., FISHER, J., MAINES-BANDIERA, S., SOMASIRI, A. & ROSKELLEY, C. D. 1999. E-cadherin induces mesenchymal-to-epithelial transition in human ovarian surface epithelium. *Proceedings of the National Academy of Sciences of the United States of America,* 96**,** 6249-6254.

AUERSPERG, N., SIEMENS, C. H. & MYRDAL, S. E. 1984. Human ovarian surface epithelium in primary culture. *In Vitro,* 20**,** 743-55.

AUERSPERG, N., WONG, A. S., CHOI, K. C., KANG, S. K. & LEUNG, P. C. 2001. Ovarian surface epithelium: biology, endocrinology, and pathology. *Endocr Rev,* 22**,** 255-88.

AUNER, V., KRIEGSHAUSER, G., TONG, D., HORVAT, R., REINTHALLER, A., MUSTEA, A. & ZEILLINGER, R. 2009. KRAS mutation analysis in ovarian samples using a high sensitivity biochip assay. *BMC Cancer,* 9**,** 111.

AURANEN, A., SONG, H., WATERFALL, C., DICIOCCIO, R. A., KUSCHEL, B., KJAER, S. K., HOGDALL, E., HOGDALL, C., STRATTON, J., WHITTEMORE, A. S., EASTON, D. F., PONDER, B. A. J., NOVIK, K. L., DUNNING, A. M., GAYTHER, S. & PHAROAH, P. D. P. 2005. Polymorphisms in DNA repair genes and epithelial ovarian cancer risk. *International Journal of Cancer,* 117**,** 611-618.

AYOUB, N., JEYASEKHARAN, A. D., BERNAL, J. A. & VENKITARAMAN, A. R. 2008. HP1-[bgr] mobilization promotes chromatin changes that initiate the DNA damage response. *Nature,* 453**,** 682-686.

BADGWELL, D. B., LU, Z., LE, K., GAO, F., YANG, M., SUH, G. K., BAO, J. J., DAS, P., ANDREEFF, M., CHEN, W., YU, Y., AHMED, A. A., S-L LIAO, W. & BAST, R. C. 2011. The tumor-suppressor gene ARHI (DIRAS3) suppresses ovarian cancer cell migration through inhibition of the Stat3 and FAK/Rho signaling pathways. *Oncogene*.

BAI, L. & ZHU, W.-G. 2006. p53 : Structure , Function and Therapeutic Applications. *Online,* 2**,** 141-153.

BAKER, V. V., BORST, M. P., DIXON, D., HATCH, K. D., SHINGLETON, H. M. & MILLER, D. 1990. c-myc amplification in ovarian cancer. *Gynecol Oncol,* 38**,** 340-2.

BALAJEE, A. S. & BOHR, V. A. 2000. Genomic heterogeneity of nucleotide excision repair. *Gene,* 250**,** 15-30.

BANKHEAD, C. R., COLLINS, C., STOKES-LAMPARD, H., ROSE, P., WILSON, S.,

CLEMENTS, A., MANT, D., KEHOE, S. T. & AUSTOKER, J. 2008. Identifying symptoms of ovarian cancer: a qualitative and quantitative study. *Bjog,* 115**,** 1008-14.
BANNAI, M., HIGUCHI, K., AKESAKA, T., FURUKAWA, M., YAMAOKA, M., SATO, K. & TOKUNAGA, K. 2004. Single-nucleotide-polymorphism genotyping for whole-genome-amplified samples using automated fluorescence correlation spectroscopy. *Anal Biochem,* 327**,** 215-21.
BANSAL, V., LIBIGER, O., TORKAMANI, A. & SCHORK, N. J. 2010. Statistical analysis strategies for association studies involving rare variants. *Nat Rev Genet,* 11**,** 773-85.
BARRETT, J. C., FRY, B., MALLER, J. & DALY, M. J. 2005. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics,* 21**,** 263-5.
BARSOTTI AM, B. R., LAPTENKO O, HUPPI K, CAPLEN NJ, PRIVES C. 2012. p53-Dependent induction of PVT1 and miR-1204. *J Biol Chem.,* 287**,** 2509-19.
BARTON, C. A., HACKER, N. F., CLARK, S. J. & O'BRIEN, P. M. 2008. DNA methylation changes in ovarian cancer: implications for early diagnosis, prognosis and treatment. *Gynecol Oncol,* 109**,** 129-39.
BATES, R. C. & MERCURIO, A. M. 2005. The epithelial-mesenchymal transition (EMT) and colorectal cancer progression. *Cancer Biol Ther,* 4**,** 365-70.
BAUR, F., NAU, K., SADIC, D., ALLWEISS, L., ELSÄSSER, H.-P., GILLEMANS, N., DE WIT, T., KRÜGER, I., VOLLMER, M., PHILIPSEN, S. & SUSKE, G. 2010. Specificity Protein 2 (Sp2) Is Essential for Mouse Development and Autonomous Proliferation of Mouse Embryonic Fibroblasts. *PLoS ONE,* 5**,** e9587.
BECK-ENGESER GB, L. A., HUPPI K, CAPLEN NJ, WANG BB, WABL M 2008. Pvt1-encoded microRNAs in oncogenesis. . *Retrovirology,* 5.
BECKER, J., WENDLAND, J. R., HAENISCH, B., NOTHEN, M. M. & SCHUMACHER, J. 2011. A systematic eQTL study of cis-trans epistasis in 210 HapMap individuals. *Eur J Hum Genet*.
BELIVEAU, A., MOTT, J. D., LO, A., CHEN, E. I., KOLLER, A. A., YASWEN, P., MUSCHLER, J. & BISSELL, M. J. 2010. Raf-induced MMP9 disrupts tissue architecture of human breast cells in three-dimensional culture and is necessary for tumor growth in vivo. *Genes Dev,* 24**,** 2800-11.
BELL, D. A. 2005. Origins and molecular pathology of ovarian cancer. *Mod Pathol,* 18**,** S19-32.
BELL, J., PAI, A., PICKRELL, J., GAFFNEY, D., PIQUE-REGI, R., DEGNER, J., GILAD, Y. & PRITCHARD, J. 2011. DNA methylation patterns associate with genetic and gene expression variation in HapMap cell lines. *Genome Biology,* 12**,** R10.
BELLACOSA, A., DE FEO, D., GODWIN, A. K., BELL, D. W., CHENG, J. Q., ALTOMARE, D. A., WAN, M., DUBEAU, L., SCAMBIA, G., MASCIULLO, V., FERRANDINA, G., BENEDETTI PANICI, P., MANCUSO, S., NERI, G. & TESTA, J. R. 1995. Molecular alterations of the AKT2 oncogene in ovarian and breast carcinomas. *Int J Cancer,* 64**,** 280-5.
BEN DAVID, Y., CHETRIT, A., HIRSH-YECHEZKEL, G., FRIEDMAN, E., BECK, B. D., BELLER, U., BEN-BARUCH, G., FISHMAN, A., LEVAVI, H., LUBIN, F., MENCZER, J., PIURA, B., STRUEWING, J. P. & MODAN, B. 2002. Effect of BRCA mutations on the length of survival in epithelial ovarian tumors. *J Clin Oncol,* 20**,** 463-6.
BERAL, V., BULL, D., GREEN, J. & REEVES, G. 2007. Ovarian cancer and hormone replacement therapy in the Million Women Study. *Lancet,* 369**,** 1703-10.
BERCHUCK, A., IVERSEN, E., LANCASTER, J., PITTMAN, J., LUO, J., LEE, P., MURPHY, S., DRESSMAN, H., FEBBO, P. & WEST, M. 2005. Patterns of gene expression that characterize long-term survival in advanced stage serous ovarian cancers. *Clin Cancer Res,* 11**,** 3686 - 3696.
BERCHUCK, A., IVERSEN, E., LUO, J., CLARKE, J., HORNE, H., LEVINE, D., BOYD, J., ALONSO, M., SECORD, A. & BERNARDINI, M. 2009. Microarray analysis of early stage serous ovarian cancers demonstrates profiles predictive of favorable outcome. *Clin Cancer Res,* 15**,** 2448 - 2455.
BERCHUCK, A., KOHLER, M. F., MARKS, J. R., WISEMAN, R., BOYD, J. & BAST, R. C., JR. 1994. The p53 tumor suppressor gene frequently is altered in gynecologic cancers. *Am J Obstet Gynecol,* 170**,** 246-52.
BERCHUCK, A., SCHILDKRAUT, J. M., MARKS, J. R. & FUTREAL, P. A. 1999. Managing hereditary ovarian cancer risk. *Cancer,* 86**,** 2517-2524.
BERNARDINI, M., BABA, T., LEE, P., BARNETT, J., SFAKIANOS, G., SECORD, A., MURPHY, S., IVERSEN, E., MARKS, J. & BERCHUCK, A. 2010. Expression signatures of TP53

mutations in serous ovarian cancers. *BMC Cancer,* 10**,** 237.

BEROUKHIM, R., LIN, M., PARK, Y., HAO, K., ZHAO, X., GARRAWAY, L. A., FOX, E. A., HOCHBERG, E. P., MELLINGHOFF, I. K., HOFER, M. D., DESCAZEAUD, A., RUBIN, M. A., MEYERSON, M., WONG, W. H., SELLERS, W. R. & LI, C. 2006. Inferring loss-of-heterozygosity from unpaired tumors using high-density oligonucleotide SNP arrays. *PLoS Comput Biol,* 2**,** 12.

BIRD, A. P. & WOLFFE, A. P. 1999. Methylation-induced repression--belts, braces, and chromatin. *Cell,* 99**,** 451-4.

BJORNSSON, H., ALBERT, T., LADD-ACOSTA, C., GREEN, R., RONGIONE, M., MIDDLE, C., IRIZARRY, R., BROMAN, K. & FEINBERG, A. 2008. SNP-specific array-based allele-specific expression analysis. *Genome Res,* 18**,** 771 - 779.

BOKS, M. P., DERKS, E. M., WEISENBERGER, D. J., STRENGMAN, E., JANSON, E., SOMMER, I. E., KAHN, R. S. & OPHOFF, R. A. 2009. The Relationship of DNA Methylation with Age, Gender and Genotype in Twins and Healthy Controls. *PLoS ONE,* 4**,** e6767.

BOLDERSON, E., RICHARD, D. J., ZHOU, B.-B. S. & KHANNA, K. K. 2009. Recent Advances in Cancer Therapy Targeting Proteins Involved in DNA Double-Strand Break Repair. *Clinical Cancer Research,* 15**,** 6314-6320.

BOLTON, K. L., CHENEVIX-TRENCH, G., GOH, C., SADETZKI, S., RAMUS, S. J., KARLAN, B. Y., LAMBRECHTS, D., DESPIERRE, E., BARROWDALE, D., MCGUFFOG, L., HEALEY, S., EASTON, D. F., SINILNIKOVA, O., BENITEZ, J., GARCIA, M. J., NEUHAUSEN, S., GAIL, M. H., HARTGE, P., PEOCK, S., FROST, D., EVANS, D. G., EELES, R., GODWIN, A. K., DALY, M. B., KWONG, A., MA, E. S., LAZARO, C., BLANCO, I., MONTAGNA, M., D'ANDREA, E., NICOLETTO, M. O., JOHNATTY, S. E., KJAER, S. K., JENSEN, A., HOGDALL, E., GOODE, E. L., FRIDLEY, B. L., LOUD, J. T., GREENE, M. H., MAI, P. L., CHETRIT, A., LUBIN, F., HIRSH-YECHEZKEL, G., GLENDON, G., ANDRULIS, I. L., TOLAND, A. E., SENTER, L., GORE, M. E., GOURLEY, C., MICHIE, C. O., SONG, H., TYRER, J., WHITTEMORE, A. S., MCGUIRE, V., SIEH, W., KRISTOFFERSSON, U., OLSSON, H., BORG, A., LEVINE, D. A., STEELE, L., BEATTIE, M. S., CHAN, S., NUSSBAUM, R. L., MOYSICH, K. B., GROSS, J., CASS, I., WALSH, C., LI, A. J., LEUCHTER, R., GORDON, O., GARCIA-CLOSAS, M., GAYTHER, S. A., CHANOCK, S. J., ANTONIOU, A. C. & PHAROAH, P. D. 2012. Association between BRCA1 and BRCA2 mutations and survival in women with invasive epithelial ovarian cancer. *Jama,* 307**,** 382-90.

BOLTON, K. L., TYRER, J., SONG, H., RAMUS, S. J., NOTARIDOU, M., JONES, C., SHER, T., GENTRY-MAHARAJ, A., WOZNIAK, E., TSAI, Y.-Y., WEIDHAAS, J., PAIK, D., VAN DEN BERG, D. J., STRAM, D. O., PEARCE, C. L., WU, A. H., BREWSTER, W., ANTON-CULVER, H., ZIOGAS, A., NAROD, S. A., LEVINE, D. A., KAYE, S. B., BROWN, R., PAUL, J., FLANAGAN, J., SIEH, W., MCGUIRE, V., WHITTEMORE, A. S., CAMPBELL, I., GORE, M. E., LISSOWSKA, J., YANG, H. P., MEDREK, K., GRONWALD, J., LUBINSKI, J., JAKUBOWSKA, A., LE, N. D., COOK, L. S., KELEMEN, L. E., BROOK-WILSON, A., MASSUGER, L. F. A. G., KIEMENEY, L. A., ABEN, K. K. H., VAN ALTENA, A. M., HOULSTON, R., TOMLINSON, I., PALMIERI, R. T., MOORMAN, P. G., SCHILDKRAUT, J., IVERSEN, E. S., PHELAN, C., VIERKANT, R. A., CUNNINGHAM, J. M., GOODE, E. L., FRIDLEY, B. L., KRUGER-KJAER, S., BLAEKER, J., HOGDALL, E., HOGDALL, C., GROSS, J., KARLAN, B. Y., NESS, R. B., EDWARDS, R. P., ODUNSI, K., MOYISCH, K. B., BAKER, J. A., MODUGNO, F., HEIKKINENEN, T., BUTZOW, R., NEVANLINNA, H., LEMINEN, A., BOGDANOVA, N., ANTONENKOVA, N., DOERK, T., HILLEMANNS, P., DURST, M., RUNNEBAUM, I., THOMPSON, P. J., CARNEY, M. E., GOODMAN, M. T., LURIE, G., WANG-GOHRKE, S., HEIN, R., CHANG-CLAUDE, J., ROSSING, M. A., CUSHING-HAUGEN, K. L., DOHERTY, J., CHEN, C., RAFNAR, T., BESENBACHER, S., SULEM, P., STEFANSSON, K., BIRRER, M. J., TERRY, K. L., HERNANDEZ, D., CRAMER, D. W., VERGOTE, I., AMANT, F., LAMBRECHTS, D., DESPIERRE, E., et al. 2010. Common variants at 19p13 are associated with susceptibility to ovarian cancer. *Nat Genet,* 42**,** 880-884.

BONADONA, V., BONAITI, B., OLSCHWANG, S., GRANDJOUAN, S., HUIART, L., LONGY, M., GUIMBAUD, R., BUECHER, B., BIGNON, Y. J., CARON, O., COLAS, C., NOGUES, C., LEJEUNE-DUMOULIN, S., OLIVIER-FAIVRE, L., POLYCARPE-OSAER, F., NGUYEN, T. D., DESSEIGNE, F., SAURIN, J. C., BERTHET, P., LEROUX, D.,

DUFFOUR, J., MANOUVRIER, S., FREBOURG, T., SOBOL, H., LASSET, C. & BONAITI-PELLIE, C. 2011. Cancer risks associated with germline mutations in MLH1, MSH2, and MSH6 genes in Lynch syndrome. *Jama,* 305**,** 2304-10.

BONNEAU, D. & LONGY, M. 2000. Mutations of the human PTEN gene. *Hum Mutat,* 16**,** 109-22.

BOS, J. 1989. ras oncogenes in human cancer: a review. *Cancer Res,* 49**,** 4682 - 4689.

BOWEN, N. J., LOGANI, S., DICKERSON, E. B., KAPA, L. B., AKHTAR, M., BENIGNO, B. B. & MCDONALD, J. F. 2007. Emerging roles for PAX8 in ovarian cancer and endosalpingeal development. *Gynecol Oncol,* 104**,** 331-7.

BOYD, J., SONODA, Y., FEDERICI, M. G., BOGOMOLNIY, F., RHEI, E., MARESCO, D. L., SAIGO, P. E., ALMADRONES, L. A., BARAKAT, R. R., BROWN, C. L., CHI, D. S., CURTIN, J. P., POYNOR, E. A. & HOSKINS, W. J. 2000. Clinicopathologic features of BRCA-linked and sporadic ovarian cancer. *Jama,* 283**,** 2260-5.

BRACHNER, A., BRAUN, J., GHODGAONKAR, M., CASTOR, D., ZLOPASA, L., EHRLICH, V., JIRICNY, J., GOTZMANN, J., KNASMÜLLER, S. & FOISNER, R. 2012. The endonuclease Ankle1 requires its LEM and GIY-YIG motifs for DNA cleavage in vivo. *Journal of Cell Science*.

BRADSHAW, D. M. & ARCECI, R. J. 1998. Clinical relevance of transmembrane drug efflux as a mechanism of multidrug resistance. *J Clin Oncol,* 16**,** 3674-90.

BRAY, N., BUCKLAND, P., OWEN, M. & O'DONOVAN, M. 2003. Cis-acting variation in the expression of a high proportion of genes in human brain. *Hum Genet,* 113**,** 149 - 153.

BREST, P., LAPAQUETTE, P., SOUIDI, M., LEBRIGAND, K., CESARO, A., VOURET-CRAVIARI, V., MARI, B., BARBRY, P., MOSNIER, J.-F., HEBUTERNE, X., HAREL-BELLAN, A., MOGRABI, B., DARFEUILLE-MICHAUD, A. & HOFMAN, P. 2011. A synonymous variant in IRGM alters a binding site for miR-196 and causes deregulation of IRGM-dependent xenophagy in Crohn's disease. *Nat Genet,* 43**,** 242-245.

BROZEK, I., OCHMAN, K., DEBNIAK, J., MORZUCH, L., RATAJSKA, M., STEPNOWSKA, M., STUKAN, M., EMERICH, J. & LIMON, J. 2009. Loss of heterozygosity at BRCA1/2 loci in hereditary and sporadic ovarian cancers. *J Appl Genet,* 50**,** 379-84.

BRUN, J. L., FEYLER, A., CHENE, G., SAUREL, J., BRUN, G. & HOCKE, C. 2000. Long-term results and prognostic factors in patients with epithelial ovarian cancer. *Gynecol Oncol,* 78**,** 21-7.

BRUN, J.-L., FEYLER, A., CHÊNE, G., SAUREL, J., BRUN, G. & HOCKÉ, C. 2000. Long-Term Results and Prognostic Factors in Patients with Epithelial Ovarian Cancer. *Gynecologic Oncology,* 78**,** 21-27.

BULLER, R., LALLAS, T., SHAHIN, M., SOOD, A., HATTERMAN-ZOGG, M., ANDERSON, B., SOROSKY, J. & KIRBY, P. 2001. The p53 mutational spectrum associated with BRCA1 mutant ovarian cancer. *Clin Cancer Res,* 7**,** 831 - 838.

BURBACH BJ, S. R., INGRAM MA, MITCHELL JS, SHIMIZU Y. 2011. The pleckstrin homology domain in the SKAP55 adapter protein defines the ability of the adapter protein ADAP to regulate integrin function and NF-kappaB activation. *J Immunol.,* 186**,** 6227-37.

BURBEE, D. G., FORGACS, E., ZOCHBAUER-MULLER, S., SHIVAKUMAR, L., FONG, K., GAO, B., RANDLE, D., KONDO, M., VIRMANI, A., BADER, S., SEKIDO, Y., LATIF, F., MILCHGRUB, S., TOYOOKA, S., GAZDAR, A. F., LERMAN, M. I., ZABAROVSKY, E., WHITE, M. & MINNA, J. D. 2001. Epigenetic inactivation of RASSF1A in lung and breast cancers and malignant phenotype suppression. *J Natl Cancer Inst,* 93**,** 691-9.

BURDETTE, J. E., OLIVER, R. M., ULYANOV, V., KILEN, S. M., MAYO, K. E. & WOODRUFF, T. K. 2007. Ovarian Epithelial Inclusion Cysts in Chronically Superovulated CD1 and Smad2 Dominant-Negative Mice. *Endocrinology,* 148**,** 3595-3604.

BURTON, E. R., GAFFAR, A., LEE, S. J., ADESHUKO, F., WHITNEY, K. D., CHUNG, J.-Y., HEWITT, S. M., HUANG, G. S., GOLDBERG, G. L., LIBUTTI, S. K. & KWON, M. 2011. Downregulation of Filamin A Interacting Protein 1-Like is Associated with Promoter Methylation and Induces an Invasive Phenotype in Ovarian Cancer. *Molecular Cancer Research*.

BURTT, N. P. 2011. Whole-Genome Amplification Using Φ29 DNA Polymerase. *Cold Spring Harbor Protocols,* 2011**,** pdb.prot5552.

BYSKOV, A. G. 1986. Differentiation of mammalian embryonic gonad. *Physiol Rev,* 66**,** 71-117.

CADUFF, R., SVOBODA-NEWMAN, S., FERGUSON, A., JOHNSTON, C. & FRANK, T. 1999. Comparison of mutations of Ki-RAS and p53 immunoreactivity in borderline and malignant epithelial ovarian tumors. *Am J Surg Pathol,* 23**,** 323 - 328.

CALIN, G. A., LIU, C.-G., FERRACIN, M., HYSLOP, T., SPIZZO, R., SEVIGNANI, C., FABBRI, M., CIMMINO, A., LEE, E. J., WOJCIK, S. E., SHIMIZU, M., TILI, E., ROSSI, S., TACCIOLI, C., PICHIORRI, F., LIU, X., ZUPO, S., HERLEA, V., GRAMANTIERI, L., LANZA, G., ALDER, H., RASSENTI, L., VOLINIA, S., SCHMITTGEN, THOMAS D., KIPPS, T. J., NEGRINI, M. & CROCE, C. M. 2007. Ultraconserved Regions Encoding ncRNAs Are Altered in Human Leukemias and Carcinomas. *Cancer Cell,* 12**,** 215-229.

CALLAHAN, M. J., CRUM, C. P., MEDEIROS, F., KINDELBERGER, D. W., ELVIN, J. A., GARBER, J. E., FELTMATE, C. M., BERKOWITZ, R. S. & MUTO, M. G. 2007. Primary fallopian tube malignancies in BRCA-positive women undergoing surgery for ovarian cancer risk reduction. *J Clin Oncol,* 25**,** 3985-90.

CALZA, S., HALL, P., AUER, G., BJOHLE, J., KLAAR, S., KRONENWETT, U., LIU, E., MILLER, L., PLONER, A. & SMEDS, J. 2006. Intrinsic molecular signature of breast cancer in a population-based cohort of 412 patients. *Breast Cancer Res,* 8**,** R34.

CAMPBELL, C., KIRBY, A., NEMESH, J., DALY, M. & HIRSCHHORN, J. 2008. A survey of allelic imbalance in F1 mice. *Genome Res,* 18**,** 555 - 563.

CAMPBELL, I. G., RUSSELL, S. E., CHOONG, D. Y., MONTGOMERY, K. G., CIAVARELLA, M. L., HOOI, C. S., CRISTIANO, B. E., PEARSON, R. B. & PHILLIPS, W. A. 2004. Mutation of the PIK3CA gene in ovarian and breast cancer. *Cancer Res,* 64**,** 7678-81.

CANNISTRA, S. A. 1997. *BRCA1 mutations and survival in women with ovarian cancer*, N Engl J Med. 1997 Apr 24;336(17):1254; author reply 1256-7.

CANNISTRA, S. A. 2004. Cancer of the Ovary. *New England Journal of Medicine,* 351**,** 2519-2529.

CANTILE, M., FRANCO, R., TSCHAN, A., BAUMHOER, D. & ZLOBEC, I. 2009. HOX D13 expression across 79 tumor tissue types. *International Journal of Cancer,* 125**,** 1532-1541.

CANZIAN, F., SALOVAARA, R., HEMMINKI, A., KRISTO, P., CHADWICK, R. B., AALTONEN, L. A. & DE LA CHAPELLE, A. 1996. Semiautomated assessment of loss of heterozygosity and replication error in tumors. *Cancer Res,* 56**,** 3331-7.

CARLSON, J., MIRON, A., JARBOE, E., PARAST, M., HIRSCH, M., LEE, Y., MUTO, M., KINDELBERGER, D. & CRUM, C. 2008. Serous tubal intraepithelial carcinoma: its potential role in primary peritoneal serous carcinoma and serous cancer prevention. *J Clin Oncol,* 26**,** 4160 - 4165.

CARRAMUSA, L., CONTINO, F., FERRO, A., MINAFRA, L. & PERCONTI, G. 2007. The PVT-1 oncogene is a myc protein target is overexpressed in transformed cells. *Journal of cellular physiology,* 213**,** 511-518.

CARRIO, M., ARDERIU, G., MYERS, C. & BOUDREAU, N. J. 2005. Homeobox D10 Induces Phenotypic Reversion of Breast Tumor Cells in a Three-Dimensional Culture Model. *Cancer Research,* 65**,** 7177-7185.

CATTORETTI, G., RILKE, F., ANDREOLA, S., DAMATO, L. & DELIA, D. 1988. P53 expression in breast cancer. *Int J Cancer,* 41**,** 178 - 183.

CHAN, K. Y., OZCELIK, H., CHEUNG, A. N., NGAN, H. Y. & KHOO, U. S. 2002. Epigenetic factors controlling the BRCA1 and BRCA2 genes in sporadic ovarian cancer. *Cancer Res,* 62**,** 4151-6.

CHANDLER, L. A., GHAZI, H., JONES, P. A., BOUKAMP, P. & FUSENIG, N. E. 1987. Allele-specific methylation of the human c-Ha-ras-1 gene. *Cell,* 50**,** 711-717.

CHANG, C. C., YANG, S. H., CHIEN, C. C., CHEN, S. H., PAN, S., LEE, C. L., LIN, C. M., SUN, H. L., HUANG, C. C., WU, Y. Y., YANG, R. N. & HUANG, C. J. 2009. Clinical meaning of age-related expression of fecal cytokeratin 19 in colorectal malignancy. *BMC Cancer,* 9**,** 376.

CHEN, E. Y., MEHRA, K., MEHRAD, M., NING, G., MIRON, A., MUTTER, G. L., MONTE, N., QUADE, B. J., MCKEON, F. D., YASSIN, Y., XIAN, W. & CRUM, C. P. 2010. Secretory cell outgrowth, PAX2 and serous carcinogenesis in the Fallopian tube. *J Pathol,* 222**,** 110-6.

CHEN, L., SU, L., LI, J., ZHENG, Y., YU, B., YU, Y., YAN, M., GU, Q., ZHU, Z. & LIU, B. 2012. Hypermethylated FAM5C and MYLK in serum as diagnosis and pre-warning markers for gastric cancer. *Dis Markers,* 32**,** 195-202.

CHEN, P. L., CHEN, C. F., CHEN, Y., XIAO, J., SHARP, Z. D. & LEE, W. H. 1998. The BRC repeats in BRCA2 are critical for RAD51 binding and resistance to methyl methanesulfonate treatment. *Proc Natl Acad Sci U S A,* 95**,** 5287-92.

CHEN, S., IVERSEN, E. S., FRIEBEL, T., FINKELSTEIN, D., WEBER, B. L., EISEN, A.,

PETERSON, L. E., SCHILDKRAUT, J. M., ISAACS, C., PESHKIN, B. N., CORIO, C., LEONDARIDIS, L., TOMLINSON, G., DUTSON, D., KERBER, R., AMOS, C. I., STRONG, L. C., BERRY, D. A., EUHUS, D. M. & PARMIGIANI, G. 2006. Characterization of BRCA1 and BRCA2 mutations in a large United States sample. *J Clin Oncol,* 24**,** 863-71.

CHEN, X., WEAVER, J., BOVE, B., VANDERVEER, L., WEIL, S., MIRON, A., DALY, M. & GODWIN, A. 2008. Allelic Imbalance in BRCA1 and BRCA2 Gene Expression Is Associated with an Increased Breast Cancer Risk. *Hum Mol Genet,* 17**,** 1336 - 1348.

CHENE, G., PENAULT-LLORCA, F., LE BOUËDEC, G., MISHELLANY, F., DAUPLAT, M. M., JAFFEUX, P., AUBLET-CUVELIER, B., POULY, J. L., DECHELOTTE, P. & DAUPLAT, J. 2009. Ovarian epithelial dysplasia after ovulation induction: time and dose effects. *Hum Reprod,* 24**,** 132-138.

CHENEVIX-TRENCH, G., LEARY, J., KERR, J., MICHEL, J., KEFFORD, R., HURST, T., PARSONS, P. G., FRIEDLANDER, M. & KHOO, S. K. 1992. Frequent loss of heterozygosity on chromosome 18 in ovarian adenocarcinoma which does not always include the DCC locus. *Oncogene,* 7**,** 1059-65.

CHENG, W., LIU, J., YOSHIDA, H., ROSEN, D. & NAORA, H. 2005. Lineage infidelity of epithelial ovarian cancers is controlled by HOX genes that specify regional identity in the reproductive tract. *Nat Med,* 11**,** 531-537.

CHETRIT, A., HIRSH-YECHEZKEL, G., BEN-DAVID, Y., LUBIN, F., FRIEDMAN, E. & SADETZKI, S. 2008. Effect of BRCA1/2 mutations on long-term survival of patients with invasive ovarian cancer: the national Israeli study of ovarian cancer. *J Clin Oncol,* 26**,** 20-5.

CHEUNG, V., BRUZEL, A., BURDICK, J., MORLEY, M., DEVLIN, J. & SPIELMAN, R. 2008. Monozygotic Twins Reveal Germline Contribution to Allelic Expression Differences. *Am J Hum Genet,* 82**,** 1357 - 1360.

CHEUNG, V., CONLIN, L., WEBER, T., ARCARO, M., JEN, K., MORLEY, M. & SPIELMAN, R. 2003. Natural variation in human gene expression assessed in lymphoblastoid cells. *Nat Genet,* 33**,** 422 - 425.

CHEUNG, V., SPIELMAN, R., EWENS, K., WEBER, T., MORLEY, M. & BURDICK, J. 2005. Mapping determinants of human gene expression by regional and genome-wide association. *Nature,* 437**,** 1365 - 1369.

CHEUNG, V. G. & NELSON, S. F. 1996. Whole genome amplification using a degenerate oligonucleotide primer allows hundreds of genotypes to be performed on less than one nanogram of genomic DNA. *Proc Natl Acad Sci U S A,* 93**,** 14676-9.

CHEVILLE, J. C., RAO, S., ICZKOWSKI, K. A., LOHSE, C. M. & PANKRATZ, V. S. 2000. Cytokeratin expression in seminoma of the human testis. *Am J Clin Pathol,* 113**,** 583-8.

CHINNADURAI, G. 2006. CtIP, a candidate tumor susceptibility gene is a team player with luminaries. *Biochim Biophys Acta,* 1**,** 67-73.

CHOCK, K. L., ALLISON, J. M., SHIMIZU, Y. & ELSHAMY, W. M. 2010. BRCA1-IRIS overexpression promotes cisplatin resistance in ovarian cancer cells. *Cancer Res,* 70**,** 8782-91.

CHOY, E., YELENSKY, R., BONAKDAR, S., PLENGE, R., SAXENA, R., DE JAGER, P., SHAW, S., WOLFISH, C., SLAVIK, J., COTSAPAS, C., RIVAS, M., DERMITZAKIS, E., CAHIR-MCFARLAND, E., KIEFF, E., HAFLER, D., DALY, M. & ALTSHULER, D. 2008. Genetic analysis of human traits in vitro: drug response and gene expression in lymphoblastoid cell lines. *PLoS Genet,* 4**,** e1000287.

CHRISTMANN, M., TOMICIC, M. T., ROOS, W. P. & KAINA, B. 2003. Mechanisms of human DNA repair: an update. *Toxicology,* 193**,** 3-34.

CIBULA, D., WIDSCHWENDTER, M., ZIKAN, M. & DUSEK, L. 2011. Underlying mechanisms of ovarian cancer risk reduction after tubal ligation. *Acta Obstetricia et Gynecologica Scandinavica,* 90**,** 559-563.

CLARK, S. J. & MELKI, J. 2002. DNA methylation and gene silencing in cancer: which is the guilty party? *Oncogene,* 21**,** 5380-7.

CLIBY, W., RITLAND, S., HARTMANN, L., DODSON, M., HALLING, K. C., KEENEY, G., PODRATZ, K. C. & JENKINS, R. B. 1993. Human epithelial ovarian cancer allelotype. *Cancer Res,* 53**,** 2393-8.

COLLINS, A., LONJOU, C. & MORTON, N. E. 1999. Genetic epidemiology of single-nucleotide polymorphisms. *Proc Natl Acad Sci U S A,* 96**,** 15173-7.

COMER, M. T., LEESE, H. J. & SOUTHGATE, J. 1998. Induction of a differentiated ciliated cell

phenotype in primary cultures of Fallopian tube epithelium. *Hum Reprod,* 13**,** 3114-20.
CONNOLLY, D. C., BAO, R., NIKITIN, A. Y., STEPHENS, K. C., POOLE, T. W., HUA, X., HARRIS, S. S., VANDERHYDEN, B. C. & HAMILTON, T. C. 2003. Female mice chimeric for expression of the simian virus 40 TAg under control of the MISIIR promoter develop epithelial ovarian cancer. *Cancer Res,* 63**,** 1389-97.
COOKE, I. E., SHELLING, A. N., LE MEUTH, V. G., CHARNOCK, M. L. & GANESAN, T. S. 1996. Allele loss on chromosome arm 6q and fine mapping of the region at 6q27 in epithelial ovarian cancer. *Genes Chromosomes Cancer,* 15**,** 223-33.
COOPER, G. M. & SHENDURE, J. 2011. Needles in stacks of needles: finding disease-causal variants in a wealth of genomic data. *Nat Rev Genet,* 12**,** 628-40.
CORDELL, H. J. & CLAYTON, D. G. 2005. Genetic association studies. *Lancet,* 366**,** 1121-31.
CRAMER, D. W. 1999. *Perineal talc exposure and subsequent epithelial ovarian cancer: a case-control study*, Obstet Gynecol. 1999 Jul;94(1):160-1.
CRAMER, D. W., LIBERMAN, R. F., TITUS-ERNSTOFF, L., WELCH, W. R., GREENBERG, E. R., BARON, J. A. & HARLOW, B. L. 1999. Genital talc exposure and risk of ovarian cancer. *Int J Cancer,* 81**,** 351-6.
CRITCHLOW, S. E. & JACKSON, S. P. 1998. DNA end-joining: from yeast to man. *Trends Biochem Sci,* 23**,** 394-8.
CRUM, C., DRAPKIN, R., MIRON, A., INCE, T., MUTO, M., KINDELBERGER, D. & LEE, Y. 2007. The distal fallopian tube: a new model for pelvic serous carcinogenesis. *Curr Opin Obstet Gynecol,* 19**,** 3 - 9.
CUATRECASAS, M., VILLANUEVA, A., MATIAS-GUIU, X. & PRAT, J. 1997. K-ras mutations in mucinous ovarian tumors: a clinicopathologic and molecular study of 95 cases. *Cancer,* 79**,** 1581-6.
DAFOU, D., RAMUS, S. J., CHOI, K., GRUN, B., TROTT, D. A., NEWBOLD, R. F., JACOBS, I. J., JONES, C. & GAYTHER, S. A. 2009. Chromosomes 6 and 18 induce neoplastic suppression in epithelial ovarian cancer cells. *Int J Cancer,* 124**,** 1037-44.
DANFORTH, K. N., TWOROGER, S. S., HECHT, J. L., ROSNER, B. A., COLDITZ, G. A. & HANKINSON, S. E. 2007. Breastfeeding and risk of ovarian cancer in two prospective cohorts. *Cancer Causes Control,* 18**,** 517-23.
DARCY, K. M. & BIRRER, M. J. 2010. Translational research in the Gynecologic Oncology Group: evaluation of ovarian cancer markers, profiles, and novel therapies. *Gynecol Oncol,* 117**,** 429-39.
DARCY, K. M., BRADY, W. E., BLANCATO, J. K., DICKSON, R. B., HOSKINS, W. J., MCGUIRE, W. P. & BIRRER, M. J. 2009. Prognostic relevance of c-MYC gene amplification and polysomy for chromosome 8 in suboptimally-resected, advanced stage epithelial ovarian cancers: a Gynecologic Oncology Group study. *Gynecol Oncol,* 114**,** 472-9.
DAVYDOV, E. V., GOODE, D. L., SIROTA, M., COOPER, G. M., SIDOW, A. & BATZOGLOU, S. 2010. Identifying a High Fraction of the Human Genome to be under Selective Constraint Using GERP++. *PLoS Comput Biol,* 6**,** e1001025.
DEAN, F. B., NELSON, J. R., GIESLER, T. L. & LASKEN, R. S. 2001. Rapid amplification of plasmid and phage DNA using Phi 29 DNA polymerase and multiply-primed rolling circle amplification. *Genome Res,* 11**,** 1095-9.
DELIGDISCH, L., GIL, J., KERNER, H., WU, H. S., BECK, D. & GERSHONI-BARUCH, R. 1999. Ovarian dysplasia in prophylactic oophorectomy specimens. *Cancer,* 86**,** 1544-1550.
DENG, C.-X. BRCA1: cell cycle checkpoint, genetic instability, DNA damage response and cancer evolution. *Nucleic Acids Research,* 34**,** 1416-1426.
DERMITZAKIS, E. & STRANGER, B. 2006. Genetic variation in human gene expression. *Mamm Genome,* 17**,** 503 - 508.
DESPIERRE, E., LAMBRECHTS, D., NEVEN, P., AMANT, F., LAMBRECHTS, S. & VERGOTE, I. 2010. The molecular genetic basis of ovarian cancer and its roadmap towards a better treatment. *Gynecol Oncol,* 117**,** 358-65.
DIANI-MOORE, S., RAM, P., LI, X., MONDAL, P., YOUN, D. Y., SAUVE, A. A. & RIFKIND, A. B. 2010. Identification of the Aryl Hydrocarbon Receptor Target Gene TiPARP as a Mediator of Suppression of Hepatic Gluconeogenesis by 2,3,7,8-Tetrachlorodibenzo-p-dioxin and of Nicotinamide as a Corrective Agent for This Effect. *Journal of Biological Chemistry,* 285**,** 38801-38810.
DIEBOLD, J., SEEMULLER, F. & LOHRS, U. 2003. K-RAS mutations in ovarian and

extraovarian lesions of serous tumors of borderline malignancy. *Lab Invest,* 83**,** 251 - 258.

DIETMAIER, W., HARTMANN, A., WALLINGER, S., HEINMOLLER, E., KERNER, T., ENDL, E., JAUCH, K. W., HOFSTADTER, F. & RUSCHOFF, J. 1999. Multiple mutation analyses in single tumor cells with improved whole genome amplification. *Am J Pathol,* 154**,** 83-95.

DOHERTY, A. M. & FISHER, E. M. 2003. Microcell-mediated chromosome transfer (MMCT): small cells with huge potential. *Mamm Genome,* 14**,** 583-92.

DONG, Q., KIRLEY, S., RUEDA, B., ZHAO, C., ZUKERBERG, L. & OLIVA, E. 2003. Loss of Cables, a Novel Gene on Chromosome 18q, in Ovarian Cancer. *Mod Pathol,* 16**,** 863-868.

DREW, Y. & CALVERT, H. 2008. The potential of PARP inhibitors in genetic breast and ovarian cancers. *Ann N Y Acad Sci***,** 136-45.

DU, H. & TAYLOR, H. S. 2004. Molecular Regulation of Müllerian Development by Hox Genes. *Annals of the New York Academy of Sciences,* 1034**,** 152-165.

DUBEAU, L. 2008. The cell of origin of ovarian epithelial tumours. *Lancet Oncol,* 9**,** 1191-7.

DUCASSE, M. & BROWN, M. 2006. Epigenetic aberrations and cancer. *Molecular Cancer,* 5**,** 60.

DUNN, B. K. 2003. Hypomethylation: one side of a larger picture. *Ann N Y Acad Sci,* 983**,** 28-42.

E., E. 2011. The role of Hox proteins in leukemogenesis: insights into key regulatory events in hematopoiesis. *Crit Rev Oncog,* 16**,** 65-76.

EASTON, D., POOLEY, K., DUNNING, A., PHAROAH, P., THOMPSON, D., BALLINGER, D., STRUEWING, J., MORRISON, J., FIELD, H., LUBEN, R., WAREHAM, N., AHMED, S., HEALEY, C., BOWMAN, R., COLLABORATORS, S., MEYER, K., HAIMAN, C., KOLONEL, L., HENDERSON, B., LE MARCHAND, L., BRENNAN, P., SANGRAJRANG, S., GABORIEAU, V., ODEFREY, F., SHEN, C., WU, P., WANG, H., ECCLES, D., EVANS, D. & PETO, J. 2007. Genome-wide association study identifies novel breast cancer susceptibility loci. *Nature,* 447**,** 1087 - 1093.

EASTON, D. F., BISHOP, D. T., FORD, D. & CROCKFORD, G. P. 1993. Genetic linkage analysis in familial breast and ovarian cancer: results from 214 families. The Breast Cancer Linkage Consortium. *Am J Hum Genet,* 52**,** 678-701.

EASTON, D. F., POOLEY, K. A., DUNNING, A. M., PHAROAH, P. D., THOMPSON, D., BALLINGER, D. G., STRUEWING, J. P., MORRISON, J., FIELD, H., LUBEN, R., WAREHAM, N., AHMED, S., HEALEY, C. S., BOWMAN, R., MEYER, K. B., HAIMAN, C. A., KOLONEL, L. K., HENDERSON, B. E., LE MARCHAND, L., BRENNAN, P., SANGRAJRANG, S., GABORIEAU, V., ODEFREY, F., SHEN, C. Y., WU, P. E., WANG, H. C., ECCLES, D., EVANS, D. G., PETO, J., FLETCHER, O., JOHNSON, N., SEAL, S., STRATTON, M. R., RAHMAN, N., CHENEVIX-TRENCH, G., BOJESEN, S. E., NORDESTGAARD, B. G., AXELSSON, C. K., GARCIA-CLOSAS, M., BRINTON, L., CHANOCK, S., LISSOWSKA, J., PEPLONSKA, B., NEVANLINNA, H., FAGERHOLM, R., EEROLA, H., KANG, D., YOO, K. Y., NOH, D. Y., AHN, S. H., HUNTER, D. J., HANKINSON, S. E., COX, D. G., HALL, P., WEDREN, S., LIU, J., LOW, Y. L., BOGDANOVA, N., SCHURMANN, P., DORK, T., TOLLENAAR, R. A., JACOBI, C. E., DEVILEE, P., KLIJN, J. G., SIGURDSON, A. J., DOODY, M. M., ALEXANDER, B. H., ZHANG, J., COX, A., BROCK, I. W., MACPHERSON, G., REED, M. W., COUCH, F. J., GOODE, E. L., OLSON, J. E., MEIJERS-HEIJBOER, H., VAN DEN OUWELAND, A., UITTERLINDEN, A., RIVADENEIRA, F., MILNE, R. L., RIBAS, G., GONZALEZ-NEIRA, A., BENITEZ, J., HOPPER, J. L., MCCREDIE, M., SOUTHEY, M., GILES, G. G., SCHROEN, C., JUSTENHOVEN, C., BRAUCH, H., HAMANN, U., KO, Y. D., SPURDLE, A. B., BEESLEY, J., CHEN, X., MANNERMAA, A., KOSMA, V. M., KATAJA, V., HARTIKAINEN, J., DAY, N. E., et al. 2007. Genome-wide association study identifies novel breast cancer susceptibility loci. *Nature,* 447**,** 1087-93.

EBERHARD, D., JOHNSON, B., AMLER, L., GODDARD, A., HELDENS, S., HERBST, R., INCE, W., JANNE, P., JANUARIO, T. & JOHNSON, D. 2005. Mutations in the epidermal growth factor receptor and in KRAS are predictive and prognostic indicators in patients with non-small-cell lung cancer treated with chemotherapy alone and in combination with erlotinib. *J Clin Oncol,* 23**,** 5900 - 5909.

EDMONDSON, R. J., MONAGHAN, J. M. & DAVIES, B. R. 2006. Gonadotropins mediate DNA synthesis and protection from spontaneous cell death in human ovarian surface

epithelium. *International Journal of Gynecological Cancer,* 16**,** 171-177.
ENDO, K., SASAKI, H., YANO, M., KOBAYASHI, Y., YUKIUE, H., HANEDA, H., SUZUKI, E., KAWANO, O. & FUJII, Y. 2006. Evaluation of the epidermal growth factor receptor gene mutation and copy number in non-small cell lung cancer with gefitinib therapy. *Oncol Rep,* 16**,** 533 - 541.
ENOMOTO, T., WEGHORST, C., INOUE, M., TANIZAWA, O. & RICE, J. 1991. K-ras activation occurs frequently in mucinous adenocarcinomas and rarely in other common epithelial tumors of the human ovary. *Am J Pathol,* 139**,** 777 - 785.
ERITJA, N., LLOBET, D., DOMINGO, M., SANTACANA, M., YERAMIAN, A., MATIAS-GUIU, X. & DOLCET, X. 2010. A Novel Three-Dimensional Culture System of Polarized Epithelial Cells to Study Endometrial Carcinogenesis. *Am J Pathol,* 176**,** 2722-2731.
ETEMADMOGHADAM, D., DEFAZIO, A., BEROUKHIM, R., MERMEL, C., GEORGE, J., GETZ, G., TOTHILL, R., OKAMOTO, A., RAEDER, M. B., HARNETT, P., LADE, S., AKSLEN, L. A., TINKER, A. V., LOCANDRO, B., ALSOP, K., CHIEW, Y. E., TRAFICANTE, N., FEREDAY, S., JOHNSON, D., FOX, S., SELLERS, W., URASHIMA, M., SALVESEN, H. B., MEYERSON, M. & BOWTELL, D. 2009. Integrated genome-wide DNA copy number and expression analysis identifies distinct mechanisms of primary chemoresistance in ovarian carcinomas. *Clin Cancer Res,* 15**,** 1417-27.
FABBRO, M., SAVAGE, K., HOBSON, K., DEANS, A. J., POWELL, S. N., MCARTHUR, G. A. & KHANNA, K. K. 2004. BRCA1-BARD1 Complexes Are Required for p53Ser-15 Phosphorylation and a G1/S Arrest following Ionizing Radiation-induced DNA Damage. *Journal of Biological Chemistry,* 279**,** 31251-31258.
FABJANI, G., KRIEGSHAEUSER, G., SCHUETZ, A., PRIX, L. & ZEILLINGER, R. 2005. Biochip for K-ras mutation screening in ovarian cancer. *Clin Chem,* 51**,** 784 - 787.
FABJANI, G., KUCERA, E., SCHUSTER, E., MINAI-POUR, M., CZERWENKA, K., SLIUTZ, G., LEODOLTER, S., REINER, A. & ZEILLINGER, R. 2002. Genetic alterations in endometrial hyperplasia and cancer. *Cancer Lett,* 175**,** 205 - 211.
FAJAC, A., DA SILVA, J., AHOMADEGBE, J. C., RATEAU, J. G., BERNAUDIN, J. F., RIOU, G. & BENARD, J. 1996. Cisplatin-induced apoptosis and p53 gene status in a cisplatin-resistant human ovarian carcinoma cell line. *Int J Cancer,* 68**,** 67-74.
FALLOWS, S., PRICE, J., ATKINSON, R., JOHNSTON, P., HICKEY, I. & RUSSELL, S. 2001. P53 mutation does not affect prognosis in ovarian epithelial malignancies. *J Pathol,* 194**,** 68 - 75.
FANG, L., SEKI, A. & FANG, G. 2009. SKAP associates with kinetochores and promotes the metaphase-to-anaphase transition. *Cell Cycle,* 8**,** 2819-2827.
FARMER, H., MCCABE, N., LORD, C. J., TUTT, A. N. J., JOHNSON, D. A., RICHARDSON, T. B., SANTAROSA, M., DILLON, K. J., HICKSON, I., KNIGHTS, C., MARTIN, N. M. B., JACKSON, S. P., SMITH, G. C. M. & ASHWORTH, A. 2005. Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy. *Nature,* 434**,** 917-921.
FATHALLA, M. F. 1971. Incessant ovulation--a factor in ovarian neoplasia? *Lancet,* 2**,** 163.
FAZILI, Z., SUN, W., MITTELSTAEDT, S., COHEN, C. & XU, X. X. 1999. Disabled-2 inactivation is an early step in ovarian tumorigenicity. *Oncogene,* 18**,** 3104-13.
FENG, J., BI, C., CLARK, B. S., MADY, R., SHAH, P. & KOHTZ, J. D. 2006. The Evf-2 noncoding RNA is transcribed from the Dlx-5/6 ultraconserved region and functions as a Dlx-2 transcriptional coactivator. *Genes & Development,* 20**,** 1470-1484.
FENG, L., HUANG, J. & CHEN, J. 2009. MERIT40 facilitates BRCA1 localization and DNA damage repair. *Genes & Development,* 23**,** 719-728.
FINK, D., NEBEL, S., AEBI, S., ZHENG, H., CENNI, B., NEHMÉ, A., CHRISTEN, R. D. & HOWELL, S. B. 1996. The Role of DNA Mismatch Repair in Platinum Drug Resistance. *Cancer Research,* 56**,** 4881-4886.
FITCH, M., CROSS, I., TURNER, S., ADIMOOLAM, S., LIN, C., WILLIAMS, K. & FORD, J. 2003. The DDB2 nucleotide excision repair gene product p48 enhances global genomic repair in p53 deficient human fibroblasts. *DNA Repair,* 2**,** 819 - 826.
FOLKINS, A. K., JARBOE, E. A., SALEEMUDDIN, A., LEE, Y., CALLAHAN, M. J., DRAPKIN, R., GARBER, J. E., MUTO, M. G., TWOROGER, S. & CRUM, C. P. 2008. A candidate precursor to pelvic serous cancer (p53 signature) and its prevalence in ovaries and fallopian tubes from women with BRCA mutations. *Gynecol Oncol,* 109**,** 168-73.
FOLSOM, A. R., ANDERSON, J. P. & ROSS, J. A. 2004. Estrogen replacement therapy and ovarian cancer. *Epidemiology,* 15**,** 100-4.
FORD, D. & EASTON, D. F. 1995. The genetics of breast and ovarian cancer. *Br J Cancer,* 72**,**

805-12.
FORD, D., EASTON, D. F. & PETO, J. 1995. Estimates of the gene frequency of BRCA1 and its contribution to breast and ovarian cancer incidence. *Am J Hum Genet,* 57**,** 1457-62.
FOSBRINK, M., CUDRICI, C., NICULESCU, F., BADEA, T. C., DAVID, S., SHAMSUDDIN, A., SHIN, M. L. & RUS, H. 2005. Overexpression of RGC-32 in colon cancer and other tumors. *Exp Mol Pathol,* 78**,** 116-22.
FOSTER, K. A., HARRINGTON, P., KERR, J., RUSSELL, P., DICIOCCIO, R. A., SCOTT, I. V., JACOBS, I., CHENEVIX-TRENCH, G., PONDER, B. A. & GAYTHER, S. A. 1996. Somatic and germline mutations of the BRCA2 gene in sporadic ovarian cancer. *Cancer Res,* 56**,** 3622-5.
FREEDMAN, M. L., MONTEIRO, A. N., GAYTHER, S. A., COETZEE, G. A., RISCH, A., PLASS, C., CASEY, G., DE BIASI, M., CARLSON, C., DUGGAN, D., JAMES, M., LIU, P., TICHELAAR, J. W., VIKIS, H. G., YOU, M. & MILLS, I. G. 2011. Principles for the post-GWAS functional characterization of cancer risk loci. *Nat Genet,* 43**,** 513-8.
FRIEDBERG, E. C. 2001. How nucleotide excision repair protects against cancer. *Nat Rev Cancer,* 1**,** 22-33.
FU JH, X. X., PAN L, XU W. 2008. Expression of HoxB5 mRNA and their effect on lung development in premature rats with hyperoxia-induced chronic lung disease. *Zhonghua Er Ke Za Zhi,* 46**,** 540-3.
FUJITA, M., ENOMOTO, T. & MURATA, Y. 2003. Genetic alterations in ovarian carcinoma: with specific reference to histological subtypes. *Mol Cell Endocrinol,* 202**,** 97 - 99.
FURUTA, S., JIANG, X., GU, B., CHENG, E., CHEN, P. L. & LEE, W. H. 2005. Depletion of BRCA1 impairs differentiation but enhances proliferation of mammary epithelial cells. *Proc Natl Acad Sci U S A,* 102**,** 9176-81.
FUTREAL, P. A., LIU, Q. Y., SHATTUCKEIDENS, D., COCHRAN, C., HARSHMAN, K., LIU, Q., SHATTUCK EIDENS, D., TAVTIGIAN, S., BENNETT, L. M., HAUGEN STRANO, A., SWENSEN, J. & MIKI, Y. 1994. BRCA1 mutations in primary breast and ovarian carcinomas. *Science,* 266**,** 120-122.
GALM, O., HERMAN, J. G. & BAYLIN, S. B. 2006. The fundamental role of epigenetics in hematopoietic malignancies. *Blood Rev,* 20**,** 1-13.
GARNER, E. I. O., STOKES, E. E., BERKOWITZ, R. S., MOK, S. C. & CRAMER, D. W. 2002. Polymorphisms of the Estrogen-metabolizing Genes CYP17 and Catechol-O-methyltransferase and Risk of Epithelial Ovarian Cancer. *Cancer Research,* 62**,** 3058-3062.
GAYTHER, S. A. & PHAROAH, P. D. 2010. The inherited genetics of ovarian and endometrial cancer. *Curr Opin Genet Dev,* 20**,** 231-8.
GAYTHER, S. A., RUSSELL, P., HARRINGTON, P., ANTONIOU, A. C., EASTON, D. F. & PONDER, B. A. 1999. The contribution of germline BRCA1 and BRCA2 mutations to familial ovarian cancer: no evidence for other ovarian cancer-susceptibility genes. *Am J Hum Genet,* 65**,** 1021-9.
GAYTHER, S. A., SONG, H., RAMUS, S. J., KJAER, S. K., WHITTEMORE, A. S., QUAYE, L., TYRER, J., SHADFORTH, D., HOGDALL, E., HOGDALL, C., BLAEKER, J., DICIOCCIO, R., MCGUIRE, V., WEBB, P. M., BEESLEY, J., GREEN, A. C., WHITEMAN, D. C., GOODMAN, M. T., LURIE, G., CARNEY, M. E., MODUGNO, F., NESS, R. B., EDWARDS, R. P., MOYSICH, K. B., GOODE, E. L., COUCH, F. J., CUNNINGHAM, J. M., SELLERS, T. A., WU, A. H., PIKE, M. C., IVERSEN, E. S., MARKS, J. R., GARCIA-CLOSAS, M., BRINTON, L., LISSOWSKA, J., PEPLONSKA, B., EASTON, D. F., JACOBS, I., PONDER, B. A., SCHILDKRAUT, J., PEARCE, C. L., CHENEVIX-TRENCH, G., BERCHUCK, A. & PHAROAH, P. D. 2007. Tagging single nucleotide polymorphisms in cell cycle control genes and susceptibility to invasive epithelial ovarian cancer. *Cancer Res,* 67**,** 3027-35.
GEMIGNANI, M., SCHLAERTH, A., BOGOMOLNIY, F., BARAKAT, R., LIN, O., SOSLOW, R., VENKATRAMAN, E. & BOYD, J. 2003. Role of KRAS and BRAF gene mutations in mucinous ovarian carcinoma. *Gynecol Oncol,* 90**,** 378 - 381.
GEMMA, A., TAKENAKA, K., HOSOYA, Y., MATUDA, K., SEIKE, M., KURIMOTO, F., ONO, Y., UEMATSU, K., TAKEDA, Y., HIBINO, S., YOSHIMURA, A., SHIBUYA, M. & KUDOH, S. 2001. Altered expression of several genes in highly metastatic subpopulations of a human pulmonary adenocarcinoma cell line. *European journal of cancer (Oxford, England : 1990),* 37**,** 1554-1561.
GENSCHEL, J., BAZEMORE, L. R. & MODRICH, P. 2002. Human exonuclease I is required for

5' and 3' mismatch repair. *J Biol Chem,* 277**,** 13302-11.
GERTZ, J., VARLEY, K. E., REDDY, T. E., BOWLING, K. M., PAULI, F., PARKER, S. L., KUCERA, K. S., WILLARD, H. F. & MYERS, R. M. 2011. Analysis of DNA methylation in a three-generation family reveals widespread genetic influence on epigenetic regulation. *PLoS Genet,* 7**,** 11.
GHOSH, S., SPAGNOLI, G. C., MARTIN, I., PLOEGERT, S., DEMOUGIN, P., HEBERER, M. & RESCHNER, A. 2005. Three-dimensional culture of melanoma cells profoundly affects gene expression profile: a high density oligonucleotide array study. *J Cell Physiol,* 204**,** 522-31.
GOFF, B. A., MANDEL, L. S., DRESCHER, C. W., URBAN, N., GOUGH, S., SCHURMAN, K. M., PATRAS, J., MAHONY, B. S. & ANDERSEN, M. R. 2007. Development of an ovarian cancer symptom index: possibilities for earlier detection. *Cancer,* 109**,** 221-7.
GOLDBERG, G. L. & RUNOWICZ, C. D. 1992. Ovarian carcinoma of low malignant potential, infertility, and induction of ovulation--is there a link? *Am J Obstet Gynecol,* 166**,** 853-4.
GOLUB, E. I., GUPTA, R. C., HAAF, T., WOLD, M. S. & RADDING, C. M. 1998. Interaction of human rad51 recombination protein with single-stranded DNA binding protein, RPA. *Nucleic Acids Res,* 26**,** 5388-93.
GONZALEZ-DIEGO, P., LOPEZ-ABENTE, G., POLLAN, M. & RUIZ, M. 2000. Time trends in ovarian cancer mortality in Europe (1955-1993): effect of age, birth cohort and period of death. *Eur J Cancer,* 36**,** 1816-24.
GOODE, E. L., CHENEVIX-TRENCH, G., SONG, H., RAMUS, S. J., NOTARIDOU, M., LAWRENSON, K., WIDSCHWENDTER, M., VIERKANT, R. A., LARSON, M. C., KJAER, S. K., BIRRER, M. J., BERCHUCK, A., SCHILDKRAUT, J., TOMLINSON, I., KIEMENEY, L. A., COOK, L. S., GRONWALD, J., GARCIA-CLOSAS, M., GORE, M. E., CAMPBELL, I., WHITTEMORE, A. S., SUTPHEN, R., PHELAN, C., ANTON-CULVER, H., PEARCE, C. L., LAMBRECHTS, D., ROSSING, M. A., CHANG-CLAUDE, J., MOYSICH, K. B., GOODMAN, M. T., DORK, T., NEVANLINNA, H., NESS, R. B., RAFNAR, T., HOGDALL, C., HOGDALL, E., FRIDLEY, B. L., CUNNINGHAM, J. M., SIEH, W., MCGUIRE, V., GODWIN, A. K., CRAMER, D. W., HERNANDEZ, D., LEVINE, D., LU, K., IVERSEN, E. S., PALMIERI, R. T., HOULSTON, R., VAN ALTENA, A. M., ABEN, K. K. H., MASSUGER, L. F. A. G., BROOKS-WILSON, A., KELEMEN, L. E., LE, N. D., JAKUBOWSKA, A., LUBINSKI, J., MEDREK, K., STAFFORD, A., EASTON, D. F., TYRER, J., BOLTON, K. L., HARRINGTON, P., ECCLES, D., CHEN, A., MOLINA, A. N., DAVILA, B. N., ARANGO, H., TSAI, Y.-Y., CHEN, Z., RISCH, H. A., MCLAUGHLIN, J., NAROD, S. A., ZIOGAS, A., BREWSTER, W., GENTRY-MAHARAJ, A., MENON, U., WU, A. H., STRAM, D. O., PIKE, M. C., BEESLEY, J. & WEBB, P. M. 2010. A genome-wide association study identifies susceptibility loci for ovarian cancer at 2q31 and 8q24. *Nat Genet,* 42**,** 874-879.
GOODMAN, M. T., MCDUFFIE, K., GUO, C., TERADA, K. & DONLON, T. A. 2001. CYP17 Genotype and Ovarian Cancer. *Cancer Epidemiology Biomarkers & Prevention,* 10**,** 563-564.
GOSPODINOV, A., TSANEVA, I. & ANACHKOVA, B. 2009. RAD51 foci formation in response to DNA damage is modulated by TIP49. *Int J Biochem Cell Biol,* 41**,** 925-33.
GRISARU D, H. J., PRASAD M, ALBERT M, MURPHY KJ, COVENS A, MACGREGOR PF, ROSEN B. 2007. Microarray expression identification of differentially expressed genes in serous epithelial ovarian cancer compared with bulk normal ovarian tissue and ovarian surface scrapings. *Oncol Reports,* 18**,** 1347-56.
GROUP, T. I. S. M. W. 2001. A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature,* 409**,** 928-933.
GRUBER, S. B. & KOHLMANN, W. 2003. The genetics of hereditary non-polyposis colorectal cancer. *J Natl Compr Canc Netw,* 1**,** 137-44.
GUAN, Y., KUO, W.-L., STILWELL, J. L., TAKANO, H., LAPUK, A. V., FRIDLYAND, J., MAO, J.-H., YU, M., MILLER, M. A., SANTOS, J. L., KALLOGER, S. E., CARLSON, J. W., GINZINGER, D. G., CELNIKER, S. E., MILLS, G. B., HUNTSMAN, D. G. & GRAY, J. W. 2007. Amplification of PVT1 Contributes to the Pathophysiology of Ovarian and Breast Cancer. *Clinical Cancer Research,* 13**,** 5745-5755.
GUHATHAKURTA, D., XIE, T., ANAND, M., EDWARDS, S., LI, G., WANG, S. & SCHADT, E. 2006. Cis-regulatory variations: a study of SNPs around genes showing cis-linkage in segregating mouse populations. *BMC Genomics,* 7**,** 235.
HAGOPIAN, G. S., MILLS, G. B., KHOKHAR, A. R., BAST, R. C. & SIDDIK, Z. H. 1999.

Expression of p53 in Cisplatin-resistant Ovarian Cancer Cell Lines: Modulation with the Novel Platinum Analogue (1R, 2R-Diaminocyclohexane)(trans-diacetato)(dichloro)-platinum(IV). *Clinical Cancer Research,* 5**,** 655-663.

HALL, J., PAUL, J. & BROWN, R. 2004. Critical evaluation of p53 as a prognostic marker in ovarian cancer. *Expert Rev Mol Med,* 6**,** 1-20.

HALL, J. M., LEE, M. K., NEWMAN, B., MORROW, J. E., ANDERSON, L. A., HUEY, B. & KING, M. C. 1990. Linkage of early-onset familial breast cancer to chromosome 17q21. *Science,* 250**,** 1684-9.

HAMADA, J.-I., OMATSU, T., OKADA, F., FURUUCHI, K., OKUBO, Y., TAKAHASHI, Y., TADA, M., MIYAZAKI, Y. J., TANIGUCHI, Y., SHIRATO, H., MIYASAKA, K. & MORIUCHI, T. 2001. Overexpression of homeobox gene HOXD3 induces coordinate expression of metastasis-related genes in human lung cancer cells. *International Journal of Cancer,* 93**,** 516-525.

HAMILTON, T. C., CONNOLLY, D. C., NIKITIN, A. Y., GARSON, K. & VANDERHYDEN, B. C. 2003. Translational research in ovarian cancer: a must. *Int J Gynecol Cancer,* 2**,** 220-30.

HAPMAP, I. C. 2003. The International HapMap Project. *Nature,* 426**,** 789-96.

HAQ, R. & ZANKE, B. 1998. Inhibition of apoptotic signaling pathways in cancer cells as a mechanism of chemotherapy resistance. *Cancer Metastasis Rev,* 17**,** 233-9.

HARLEY, I., ROSEN, B., RISCH, H. A., SIMINOVITCH, K., BEINER, M. E., MCLAUGHLIN, J., SUN, P. & NAROD, S. A. 2008. Ovarian cancer risk is associated with a common variant in the promoter sequence of the mismatch repair gene MLH1. *Gynecol Oncol,* 109**,** 384-7.

HARRIS, R., WHITTEMORE, A. S. & ITNYRE, J. 1992. Characteristics relating to ovarian cancer risk: collaborative analysis of 12 US case-control studies. III. Epithelial tumors of low malignant potential in white women. Collaborative Ovarian Cancer Group. *Am J Epidemiol,* 136**,** 1204-11.

HARRISON, M. L., JAMESON, C. & GORE, M. E. 2008. Mucinous ovarian cancer. *Int J Gynecol Cancer,* 18**,** 209-14.

HAVERTY, P., HON, L., KAMINKER, J., CHANT, J. & ZHANG, Z. 2009. High-resolution analysis of copy number alterations and associated expression changes in ovarian tumors. *BMC Medical Genomics,* 2**,** 21.

HAVRILESKY, L., DARCY, M., HAMDAN, H., PRIORE, R., LEON, J., BELL, J. & BERCHUCK, A. 2003. Prognostic significance of p53 mutation and p53 overexpression in advanced epithelial ovarian cancer: a Gynecologic Oncology Group Study. *J Clin Oncol,* 21**,** 3814 - 3825.

HEFLER, L. A., MUSTEA, A., KÖNSGEN, D., CONCIN, N., TANNER, B., STRICK, R., HEINZE, G., GRIMM, C., SCHUSTER, E., TEMPFER, C., REINTHALLER, A. & ZEILLINGER, R. 2007. Vascular Endothelial Growth Factor Gene Polymorphisms Are Associated with Prognosis in Ovarian Cancer. *Clinical Cancer Research,* 13**,** 898-901.

HELLEMAN, J., VAN STAVEREN, I., DINJENS, W., VAN KUIJK, P., RITSTIER, K., EWING, P., VAN DER BURG, M., STOTER, G. & BERNS, E. 2006. Mismatch repair and treatment resistance in ovarian cancer. *BMC Cancer,* 6**,** 201.

HENRIKSEN, T., TANBO, T., ABYHOLM, T., OPPEDAL, B. R., CLAUSSEN, O. P. & HOVIG, T. 1990. Epithelial cells from human fallopian tube in culture. *Hum Reprod,* 5**,** 25-31.

HERMAN, J. G. & BAYLIN, S. B. 2003. Gene silencing in cancer in association with promoter hypermethylation. *N Engl J Med,* 349**,** 2042-54.

HITCHINS, M. P., RAPKINS, R. W., KWOK, C. T., SRIVASTAVA, S., WONG, J. J., KHACHIGIAN, L. M., POLLY, P., GOLDBLATT, J. & WARD, R. L. 2011. Dominantly inherited constitutional epigenetic silencing of MLH1 in a cancer-affected family is linked to a single nucleotide variant within the 5'UTR. *Cancer Cell,* 20**,** 200-13.

HO, C., KURMAN, R., DEHARI, R., WANG, T. & SHIH IE, M. 2004. Mutations of BRAF and KRAS precede the development of ovarian serous borderline tumors. *Cancer Res,* 64**,** 6915 - 6918.

HOCHBERG, Y. & BENJAMINI, Y. 1990. More powerful procedures for multiple significance testing. *Stat Med,* 9**,** 811 - 818.

HOLSCHNEIDER, C. H. & BEREK, J. S. 2000. Ovarian cancer: epidemiology, biology, and prognostic factors. *Semin Surg Oncol,* 19**,** 3-10.

HONG JH, L. J., PARK JJ, LEE NW, LEE KW, NA JY. 2010. Expression pattern of the class I homeobox genes in ovarian carcinoma. *J Gynecol Oncol.,* 21**,** 29-37.

HOUTGRAAF, J. H., VERSMISSEN, J. & VAN DER GIESSEN, W. J. 2006. A concise review of DNA damage checkpoints and repair in mammalian cells. *Cardiovasc Revasc Med,* 7**,** 165-72.

HU, X., KIM, J. A., CASTILLO, A., HUANG, M., LIU, J. & WANG, B. 2011. NBA1/MERIT40 and BRE Interaction Is Required for the Integrity of Two Distinct Deubiquitinating Enzyme BRCC36-containing Complexes. *Journal of Biological Chemistry,* 286**,** 11734-11745.

HUDDLESTON, H. G., WONG, K.-K., WELCH, W. R., BERKOWITZ, R. S. & MOK, S. C. 2005. Clinical applications of microarray technology: creatine kinase B is an up-regulated gene in epithelial ovarian cancer and shows promise as a serum marker. *Gynecologic Oncology,* 96**,** 77-83.

HUERTAS, P. & JACKSON, S. P. 2009. Human CtIP mediates cell cycle control of DNA end resection and double strand break repair. *J Biol Chem,* 284**,** 9558-65.

HUGHES P, M. D., REID Y, PARKES H, GELBER C. 2007. The costs of using unauthenticated, over-passaged cell lines: how much more data do we need?

Hughes P, Marshall D, Reid Y, Parkes H, Gelber C. *Biotechniques.,* 43**,** 577-8, 581-2.

HUNCHAREK, M. & MUSCAT, J. 2011. Perineal talc use and ovarian cancer risk: a case study of scientific standards in environmental epidemiology. *Eur J Cancer Prev,* 20**,** 501-7.

HUNG, R. J., MCKAY, J. D., GABORIEAU, V., BOFFETTA, P., HASHIBE, M., ZARIDZE, D., MUKERIA, A., SZESZENIA-DABROWSKA, N., LISSOWSKA, J., RUDNAI, P., FABIANOVA, E., MATES, D., BENCKO, V., FORETOVA, L., JANOUT, V., CHEN, C., GOODMAN, G., FIELD, J. K., LILOGLOU, T., XINARIANOS, G., CASSIDY, A., MCLAUGHLIN, J., LIU, G., NAROD, S., KROKAN, H. E., SKORPEN, F., ELVESTAD, M. B., HVEEM, K., VATTEN, L., LINSEISEN, J., CLAVEL-CHAPELON, F., VINEIS, P., BUENO-DE-MESQUITA, H. B., LUND, E., MARTINEZ, C., BINGHAM, S., RASMUSON, T., HAINAUT, P., RIBOLI, E., AHRENS, W., BENHAMOU, S., LAGIOU, P., TRICHOPOULOS, D., HOLCATOVA, I., MERLETTI, F., KJAERHEIM, K., AGUDO, A., MACFARLANE, G., TALAMINI, R., SIMONATO, L., LOWRY, R., CONWAY, D. I., ZNAOR, A., HEALY, C., ZELENIKA, D., BOLAND, A., DELEPINE, M., FOGLIO, M., LECHNER, D., MATSUDA, F., BLANCHE, H., GUT, I., HEATH, S., LATHROP, M. & BRENNAN, P. 2008. A susceptibility locus for lung cancer maps to nicotinic acetylcholine receptor subunit genes on 15q25. *Nature,* 452**,** 633-7.

HUNTER, D., KRAFT, P., JACOBS, K., COX, D., YEAGER, M., HANKINSON, S., WACHOLDER, S., WANG, Z., WELCH, R., HUTCHINSON, A., WANG, J., YU, K., CHATTERJEE, N., ORR, N., WILLETT, W., COLDITZ, G., ZIEGLER, R., BERG, C., BUYS, S., MCCARTY, C., FEIGELSON, H., CALLE, E., THUN, M., HAYES, R., TUCKER, M., GERHARD, D., FRAUMENI, J., HOOVER, R., THOMAS, G. & CHANOCK, S. 2007. A genome-wide association study identifies alleles in FGFR2 associated with risk of sporadic postmenopausal breast cancer. *Nat Genet,* 39**,** 870 - 874.

HUPPI K, P. J., WAHLBERG BM, CAPLEN NJ. 2012. The 8q24 gene desert: an oasis of non-coding transcriptional activity. *Front Genet.,* 3.

HUPPI K, V. N., RUNFOLA T, JONES TL, MACKIEWICZ M, MARTIN SE, MUSHINSKI JF, STEPHENS R, CAPLEN NJ. 2008. The identification of microRNAs in a genomically unstable region of human chromosome 8q24. *Mol Cancer Res.,* 6**,** 212-221.

HUTSON, R., RAMSDALE, J. & WELLS, M. 1995. P53 PROTEIN EXPRESSION IN PUTATIVE PRECURSOR LESIONS OF EPITHELIAL OVARIAN-CANCER. *Histopathology,* 27**,** 367-371.

HWANG, B., FORD, J., HANAWALT, P. & CHU, G. 1999. Expression of the p48 xeroderma pigmentosum gene is p53-dependent and is involved in global genomic repair. *Proc Natl Acad Sci USA,* 96**,** 424 - 428.

HWANG, B., TOERING, S., FRANCKE, U. & CHU, G. 1998. p48 Activates a UV-damaged-DNA binding factor and is defective in xeroderma pigmentosum group E cells that lack binding activity. *Mol Cell Biol,* 18**,** 4391 - 4399.

IBANEZ DE CACERES, I., BATTAGLI, C., ESTELLER, M., HERMAN, J. G., DULAIMI, E., EDELSON, M. I., BERGMAN, C., EHYA, H., EISENBERG, B. L. & CAIRNS, P. 2004. Tumor cell-specific BRCA1 and RASSF1A hypermethylation in serum, plasma, and peritoneal fluid from ovarian cancer patients. *Cancer Res,* 64**,** 6476-81.

INAOKA, H., NOSHIRO, M. & FUKUOKA, Y. Genome-wide DNA methylation analysis in tumor and normal tissues from lung squamous cell carcinoma and serous cystadenocarcinoma. Bioinformatics and Biomedicine Workshops (BIBMW), 2010 IEEE

International Conference on, 18-18 Dec. 2010 2010. 855-856.
IVARSSON, K., SUNDFELDT, K., BRÄNNSTRÖM, M. & JANSON, P. O. 2001. Production of Steroids by Human Ovarian Surface Epithelial Cells in Culture: Possible Role of Progesterone as Growth Inhibitor. *Gynecologic Oncology,* 82**,** 116-121.
IZUMI, T., WIEDERHOLD, L. R., ROY, G., ROY, R., JAISWAL, A., BHAKAT, K. K., MITRA, S. & HAZRA, T. K. 2003. Mammalian DNA base excision repair proteins: their interactions and role in repair of oxidative DNA damage. *Toxicology,* 193**,** 43-65.
JACINTO, F. V., BALLESTAR, E., ROPERO, S. & ESTELLER, M. 2007. Discovery of Epigenetically Silenced Genes by Methylated DNA Immunoprecipitation in Colon Cancer Cells. *Cancer Research,* 67**,** 11481-11486.
JANKU, F., TSIMBERIDOU, A. M., GARRIDO-LAGUNA, I., WANG, X., LUTHRA, R., HONG, D. S., NAING, A., FALCHOOK, G. S., MORONEY, J. W., PIHA-PAUL, S. A., WHELER, J. J., MOULDER, S. L., FU, S. & KURZROCK, R. 2011. PIK3CA mutations in patients with advanced cancers treated with PI3K/AKT/mTOR axis inhibitors. *Mol Cancer Ther,* 10**,** 558-65.
JAZAERI, A. A., YEE, C. J., SOTIRIOU, C., BRANTLEY, K. R., BOYD, J. & LIU, E. T. 2002. Gene expression profiles of BRCA1-linked, BRCA2-linked, and sporadic ovarian cancers. *J Natl Cancer Inst,* 94**,** 990-1000.
JEMAL, A., BRAY, F., CENTER, M. M., FERLAY, J., WARD, E. & FORMAN, D. 2011. Global cancer statistics. *CA: A Cancer Journal for Clinicians,* 61**,** 69-90.
JESCHKE, J., VAN NESTE, L., GLOCKNER, S. C., DHIR, M., CALMON, M. F., DEREGOWSKI, V., VAN CRIEKINGE, W., VLASSENBROECK, I., KOCH, A., CHAN, T. A., COPE, L., HOOKER, C. M., SCHUEBEL, K. E., GABRIELSON, E., WINTERPACHT, A., BAYLIN, S. B., HERMAN, J. G. & AHUJA, N. 2012. Biomarkers for detection and prognosis of breast cancer identified by a functional hypermethylome screen. *Epigenetics,* 7**,** 701-9.
JINAWATH, N., VASOONTARA, C., YAP, K. L., THIAVILLE, M. M., NAKAYAMA, K., WANG, T. L. & SHIH, I. M. 2009. NAC-1, a potential stem cell pluripotency factor, contributes to paclitaxel resistance in ovarian cancer through inactivating Gadd45 pathway. *Oncogene,* 28**,** 1941-8.
JOHANNSSON, O. T., RANSTAM, J., BORG, A. & OLSSON, H. 1998. Survival of BRCA1 breast and ovarian cancer patients: a population-based study from southern Sweden. *J Clin Oncol,* 16**,** 397-404.
JONES, N. A., TURNER, J., MCILWRATH, A. J., BROWN, R. & DIVE, C. 1998. Cisplatin- and paclitaxel-induced apoptosis of ovarian carcinoma cells and the relationship between bax and bak up-regulation and the functional status of p53. *Mol Pharmacol,* 53**,** 819-26.
JONES, P. L., SCHMIDHAUSER, C. & BISSELL, M. J. 1993. Regulation of gene expression and cell function by extracellular matrix. *Crit Rev Eukaryot Gene Expr,* 3**,** 137-54.
JORDAN, S., WHITEMAN, D., PURDIE, D., GREEN, A. & WEBB, P. 2006. Does smoking increase risk of ovarian cancer? A systematic review. *Gynecologic Oncology,* 103**,** 1122-1129.
JORDAN, S. J., WHITEMAN, D. C., PURDIE, D. M., GREEN, A. C. & WEBB, P. M. 2006. Does smoking increase risk of ovarian cancer? A systematic review. *Gynecol Oncol,* 103**,** 1122-9.
KANAI, M., HAMADA, J., TAKADA, M., ASANO, T., MURAKAWA, K., TAKAHASHI, Y., MURAI, T., TADA, M., MIYAMOTO, M., KONDO, S. & MORIUCHI, T. 2010. Aberrant expressions of HOX genes in colorectal and hepatocellular carcinomas. *Oncol Rep,* 23**,** 843-51.
KANG, S. K., CHOI, K. C., TAI, C. J., AUERSPERG, N. & LEUNG, P. C. 2001. Estradiol regulates gonadotropin-releasing hormone (GnRH) and its receptor gene expression and antagonizes the growth inhibitory effects of GnRH in human ovarian surface epithelial and ovarian cancer cells. *Endocrinology,* 142**,** 580-8.
KAPRANOV, P., CHENG, J., DIKE, S., NIX, D. A., DUTTAGUPTA, R., WILLINGHAM, A. T., STADLER, P. F., HERTEL, J., HACKERMÜLLER, J., HOFACKER, I. L., BELL, I., CHEUNG, E., DRENKOW, J., DUMAIS, E., PATEL, S., HELT, G., GANESH, M., GHOSH, S., PICCOLBONI, A., SEMENTCHENKO, V., TAMMANA, H. & GINGERAS, T. R. 2007. RNA Maps Reveal New RNA Classes and a Possible Function for Pervasive Transcription. *Science,* 316**,** 1484-1488.
KARAPETIS, C., KHAMBATA-FORD, S., JONKER, D., O'CALLAGHAN, C., TU, D., TEBBUTT, N., SIMES, R., CHALCHAL, H., SHAPIRO, J. & ROBITAILLE, S. 2008. K-ras mutations

and benefit from cetuximab in advanced colorectal cancer. *N Engl J Med,* 359**,** 1757 - 1765.

KARIN, S. 2003. Cell–cell adhesion in the normal ovary and ovarian tumors of epithelial origin; an exception to the rule. *Molecular and Cellular Endocrinology,* 202**,** 89-96.

KARVE, T. M., PREET, A., SNEED, R., SALAMANCA, C., LI, X., XU, J., KUMAR, D., ROSEN, E. M. & SAHA, T. 2012. BRCA1 Regulates Follistatin Function in Ovarian Cancer and Human Ovarian Surface Epithelial Cells. *PLoS ONE,* 7**,** e37697.

KATANO, K., KONDO, A., SAFAEI, R., HOLZER, A., SAMIMI, G., MISHIMA, M., KUO, Y. M., ROCHDI, M. & HOWELL, S. B. 2002. Acquisition of resistance to cisplatin is accompanied by changes in the cellular pharmacology of copper. *Cancer Res,* 62**,** 6559-65.

KELEMEN, L. E., SELLERS, T. A., SCHILDKRAUT, J. M., CUNNINGHAM, J. M., VIERKANT, R. A., PANKRATZ, V. S., FREDERICKSEN, Z. S., GADRE, M. K., RIDER, D. N., LIEBOW, M. & GOODE, E. L. 2008. Genetic variation in the one-carbon transfer pathway and ovarian cancer risk. *Cancer Res,* 68**,** 2498-506.

KELLAND, L. 2007. The resurgence of platinum-based cancer chemotherapy. *Nat Rev Cancer,* 7**,** 573-84.

KERNER, R., SABO, E., GERSHONI-BARUCH, R., BECK, D. & BEN-IZHAK, O. 2005. Expression of cell cycle regulatory proteins in ovaries prophylactically removed from Jewish Ashkenazi BRCA1 and BRCA2 mutation carriers: Correlation with histopathology. *Gynecologic Oncology,* 99**,** 367-375.

KHAMBATA-FORD, S., GARRETT, C., MEROPOL, N., BASIK, M., HARBISON, C., WU, S., WONG, T., HUANG, X., TAKIMOTO, C. & GODWIN, A. 2007. Expression of epiregulin and amphiregulin and K-ras mutation status predict disease control in metastatic colorectal cancer patients treated with cetuximab. *J Clin Oncol,* 25**,** 3230 - 3237.

KIM, D. S., LEE, J. Y., LEE, S. M., CHOI, J. E., CHO, S. & PARK, J. Y. 2011. Promoter methylation of the RGC32 gene in nonsmall cell lung cancer. *Cancer,* 117**,** 590-596.

KIM, T.-H., CHIERA, S. L., LINDER, K. E., TREMPUS, C. S., SMART, R. C. & HOROWITZ, J. M. 2010. Overexpression of Transcription Factor Sp2 Inhibits Epidermal Differentiation and Increases Susceptibility to Wound- and Carcinogen-Induced Tumorigenesis. *Cancer Research,* 70**,** 8507-8516.

KIMCHI-SARFATY, C., OH, J. M., KIM, I. W., SAUNA, Z. E., CALCAGNO, A. M., AMBUDKAR, S. V. & GOTTESMAN, M. M. 2007. A "silent" polymorphism in the MDR1 gene changes substrate specificity. *Science,* 315**,** 525-8.

KINDELBERGER, D., LEE, Y., MIRON, A., HIRSCH, M., FELTMATE, C., MEDEIROS, F., CALLAHAN, M., GARNER, E., GORDON, R. & BIRCH, C. 2007. Intraepithelial carcinoma of the fimbria and pelvic serous carcinoma: Evidence for a causal relationship. *Am J Surg Pathol,* 31**,** 161 - 169.

KITAYAMA, J., NAGAWA, H., TSUNO, N., OSADA, T., HATANO, K., SUNAMI, E., SAITO, H. & MUTO, T. 1999. Laminin mediates tethering and spreading of colon cancer cells in physiological shear flow. *Br J Cancer,* 80**,** 1927-34.

KITTINIYOM, K., MASTRONARDI, M., ROEMER, M., WELLS, W. A., GREENBERG, E. R., TITUS-ERNSTOFF, L. & NEWSHAM, I. F. 2004. Allele-specific loss of heterozygosity at the DAL-1/4.1B (EPB41L3) tumor-suppressor gene locus in the absence of mutation. *Genes Chromosomes Cancer,* 40**,** 190-203.

KLEIVI, K., LIND, G., DIEP, C., MELING, G., BRANDAL, L., NESLAND, J., MYKLEBOST, O., ROGNUM, T., GIERCKSKY, K. & SKOTHEIM, R. 2007. Gene expression profiles of primary colorectal carcinomas, liver metastases, and carcinomatoses. *Mol Cancer,* 6**,** 2.

KNOEPFLER, P. S., SYKES, D. B., PASILLAS, M. & KAMPS, M. P. 2001. HoxB8 requires its Pbx-interaction motif to block differentiation of primary myeloid progenitors and of most cell line models of myeloid differentiation. *Oncogene,* 20**,** 5440-5448.

KNUECHEL, R., KENG, P., HOFSTAEDTER, F., LANGMUIR, V., SUTHERLAND, R. M. & PENNEY, D. P. 1990. Differentiation patterns in two- and three-dimensional culture systems of human squamous carcinoma cell lines. *Am J Pathol,* 137**,** 725-36.

KOBAYASHI, A. & BEHRINGER, R. R. 2003. Developmental genetics of the female reproductive tract in mammals. *Nature reviews. Genetics,* 4**,** 969-980.

KÖBEL, M., KALLOGER, S. E., BOYD, N., MCKINNEY, S., MEHL, E., PALMER, C., LEUNG, S., BOWEN, N. J., IONESCU, D. N., RAJPUT, A., PRENTICE, L. M., MILLER, D., SANTOS, J., SWENERTON, K., GILKS, C. B. & HUNTSMAN, D. 2008. Ovarian Carcinoma Subtypes Are Different Diseases: Implications for Biomarker Studies. *PLoS*

Med, 5**,** e232.
KOLASA, K., I., REMBISZEWSKA, A., JANIEC-JANKOWSKA, -MIESZKOWSKAA, D., LEWANDOWSKA, M., A., KONOPKA, B., KUPRYJAFICZYK & J. 2006. *PTEN mutation, expression and LOH at its locus in ovarian carcinomas. Relation to TP53, K-RAS and BRCA1 mutations,* Amsterdam, PAYS-BAS, Elsevier.
KONISHI, I., KURODA, H. & MANDAI, M. 1999. Review: gonadotropins and development of ovarian cancer. *Oncology,* 2**,** 45-8.
KOTHAPALLI, R., YODER, S., MANE, S. & LOUGHRAN, T. 2002. Microarray results: how accurate are they? *BMC Bioinformatics,* 3**,** 22.
KOWALSKI, P. J., RUBIN, M. A. & KLEER, C. G. 2003. E-cadherin expression in primary carcinomas of the breast and its distant metastases. *Breast Cancer Res,* 5**,** 26.
KRON, K. J., LIU, L., PETHE, V. V., DEMETRASHVILI, N., NESBITT, M. E., TRACHTENBERG, J., OZCELIK, H., FLESHNER, N. E., BRIOLLAIS, L., VAN DER KWAST, T. H. & BAPAT, B. 2010. DNA methylation of HOXD3 as a marker of prostate cancer progression. *Lab Invest,* 90**,** 1060-1067.
KRUGLYAK, L. 1999. Prospects for whole-genome linkage disequilibrium mapping of common disease genes. *Nat Genet,* 22**,** 139-44.
KRUGLYAK, L. & NICKERSON, D. A. 2001. Variation is the spice of life. *Nat Genet,* 27**,** 234-236.
KRUZELOCK, R. P., CUEVAS, B. D., WIENER, J. R., XU, F. J., YU, Y., CABEZA-ARVELAIZ, Y., PERSHOUSE, M., LOVELL, M. M., KILLARY, A. M., MILLS, G. B. & BAST, R. C., JR. 2000. Functional evidence for an ovarian cancer tumor suppressor gene on chromosome 22 by microcell-mediated chromosome transfer. *Oncogene,* 19**,** 6277-85.
KÜHNL, A., KAISER, M., NEUMANN, M., FRANSECKY, L., HEESCH, S., RADMACHER, M., MARCUCCI, G., BLOOMFIELD, C. D., HOFMANN, W.-K., THIEL, E. & BALDUS, C. D. 2011. High expression of IGFBP2 is associated with chemoresistance in adult acute myeloid leukemia. *Leukemia Research,* 35**,** 1585-1590.
KUMAR, S., CHENG, X., KLIMASAUSKAS, S., MI, S., POSFAI, J., ROBERTS, R. J. & WILSON, G. G. 1994. The DNA (cytosine-5) methyltransferases. *Nucleic Acids Res,* 22**,** 1-10.
KURIAN, A. W., BALISE, R. R., MCGUIRE, V. & WHITTEMORE, A. S. 2005. Histologic types of epithelial ovarian cancer: have they different risk factors? *Gynecol Oncol,* 96**,** 520-30.
KURMAN, R. & SHIH IE, M. 2008. Pathogenesis of ovarian cancer: lessons from morphology and molecular biology and their clinical implications. *Int J Gynecol Pathol,* 27**,** 151 - 160.
KURMAN, R. J. & SHIH, I.-M. 2011. Molecular pathogenesis and extraovarian origin of epithelial ovarian cancer—Shifting the paradigm. *Human pathology,* 42**,** 918-931.
KWONG, J., CHAN, F. L., WONG, K. K., BIRRER, M. J., ARCHIBALD, K. M., BALKWILL, F. R., BERKOWITZ, R. S. & MOK, S. C. 2009. Inflammatory cytokine tumor necrosis factor alpha confers precancerous phenotype in an organoid model of normal human ovarian surface epithelial cells. *Neoplasia,* 11**,** 529-41.
LA VECCHIA, C. 2006. Oral contraceptives and ovarian cancer: an update, 1998-2004. *Eur J Cancer Prev,* 15**,** 117-24.
LAMARTINE J, S. M., CINTI R, HEITZMANN F, CREAVEN M, RADOMSKI N, JOST E, LENOIR GM, ROMEO G, SYLLA BS. 1997. Molecular cloning and mapping of a human cDNA (PA2G4) that encodes a protein highly homologous to the mouse cell cycle protein p38-2G4. *Cytogenet Cell Genet.,* 78**,** 31-5.
LAMBROS, M. B., FIEGLER, H., JONES, A., GORMAN, P., ROYLANCE, R. R., CARTER, N. P. & TOMLINSON, I. P. 2005. Analysis of ovarian cancer cell lines using array-based comparative genomic hybridization. *J Pathol,* 205**,** 29-40.
LATZA, U., NIEDOBITEK, G., SCHWARTING, R., NEKARDA, H. & STEIN, H. 1990. Ber-EP4: new monoclonal antibody which distinguishes epithelia from mesothelial. *J Clin Pathol,* 43**,** 213-9.
LAVARINO, C., PILOTTI, S., OGGIONNI, M., GATTI, L., PEREGO, P., BRESCIANI, G., PIEROTTI, M. A., SCAMBIA, G., FERRANDINA, G., FAGOTTI, A., MANGIONI, C., LUCCHINI, V., VECCHIONE, F., BOLIS, G., SCARFONE, G. & ZUNINO, F. 2000. p53 gene status and response to platinum/paclitaxel-based chemotherapy in advanced ovarian carcinoma. *J Clin Oncol,* 18**,** 3936-45.
LAWRENSON, K., BENJAMIN, E., TURMAINE, M., JACOBS, I., GAYTHER, S. & DAFOU, D. 2009. In vitro three-dimensional modelling of human ovarian surface epithelial cells. *Cell Proliferation,* 42**,** 385-393.

LAWRENSON, K., GRUN, B., BENJAMIN, E., JACOBS, I. J., DAFOU, D. & GAYTHER, S. A. 2010. Senescent fibroblasts promote neoplastic transformation of partially transformed ovarian epithelial cells in a three-dimensional model of early stage ovarian cancer. *Neoplasia,* 12**,** 317-25.

LAWRENSON, K., SPROUL, D., GRUN, B., NOTARIDOU, M., BENJAMIN, E., JACOBS, I. J., DAFOU, D., SIMS, A. H. & GAYTHER, S. A. 2011. Modelling genetic and clinical heterogeneity in epithelial ovarian cancers. *Carcinogenesis,* 32**,** 1540-1549.

LEE, P., GE, B., GREENWOOD, C., SINNETT, D., FORTIN, Y., BRUNET, S., FORTIN, A., TAKANE, M., SKAMENE, E., PASTINEN, T., HALLETT, M., HUDSON, T. & SLADEK, R. 2006. Mapping cis-acting regulatory variation in recombinant congenic strains. *Physiol Genomics,* 25**,** 294 - 302.

LEE, Y., MIRON, A., DRAPKIN, R., NUCCI, M., MEDEIROS, F., SALEEMUDDIN, A., GARBER, J., BIRCH, C., MOU, H. & GORDON, R. 2007. A candidate precursor to serous carcinoma that originates in the distal fallopian tube. *J Pathol,* 211**,** 26 - 35.

LEITAO, M., SOSLOW, R., BAERGEN, R., OLVERA, N., ARROYO, C. & BOYD, J. 2004. Mutation and expression of the TP53 gene in early stage epithelial ovarian carcinoma. *Gynecol Oncol,* 93**,** 301 - 306.

LEONHARDT, K., EINENKEL, J., SOHR, S., ENGELAND, K. & HORN, L. C. 2011. p53 signature and serous tubal in-situ carcinoma in cases of primary tubal and peritoneal carcinomas and serous borderline tumors of the ovary. *Int J Gynecol Pathol,* 30**,** 417-24.

LEVANON, K., CRUM, C. & DRAPKIN, R. 2008. New insights into the pathogenesis of serous ovarian cancer and its clinical impact. *J Clin Oncol,* 26**,** 5284-93.

LEVANON, K., NG, V., PIAO, H. Y., ZHANG, Y., CHANG, M. C., ROH, M. H., KINDELBERGER, D. W., HIRSCH, M. S., CRUM, C. P., MARTO, J. A. & DRAPKIN, R. 2010. Primary ex vivo cultures of human fallopian tube epithelium as a model for serous ovarian carcinogenesis. *Oncogene,* 29**,** 1103-13.

LEVESQUE, M. A., KATSAROS, D., YU, H., ZOLA, P., SISMONDI, P., GIARDINA, G. & DIAMANDIS, E. P. 1995. Mutant p53 protein overexpression is associated with poor outcome in patients with well or moderately differentiated ovarian carcinoma. *Cancer,* 75**,** 1327-38.

LI, C., LIN, M. & LIU, J. 2004. Identification of PRC1 as the p53 target gene uncovers a novel function of p53 in the regulation of cytokinesis. *Oncogene,* 23**,** 9336 - 9347.

LI, H., HUANG, C. J. & CHOO, K. B. 2002. Expression of homeobox genes in cervical cancer. *Gynecol Oncol,* 84**,** 216-21.

LI, N. F., WILBANKS, G., BALKWILL, F., JACOBS, I. J., DAFOU, D. & GAYTHER, S. A. 2004. A modified medium that significantly improves the growth of human normal ovarian surface epithelial (OSE) cells in vitro. *Lab Invest,* 84**,** 923-31.

LI, Q., ZHU, F. & CHEN, P. 2012. miR-7 and miR-218 epigenetically control tumor suppressor genes RASSF1A and Claudin-6 by targeting HoxB3 in breast cancer. *Biochemical and Biophysical Research Communications*.

LI, Y., HU, W., SHEN, D. Y., KAVANAGH, J. J. & FU, S. 2009. Azacitidine enhances sensitivity of platinum-resistant ovarian cancer cells to carboplatin through induction of apoptosis. *Am J Obstet Gynecol,* 200**,** 25.

LICHTENSTEIN, P., HOLM, N. V., VERKASALO, P. K., ILIADOU, A., KAPRIO, J., KOSKENVUO, M., PUKKALA, E., SKYTTHE, A. & HEMMINKI, K. 2000. Environmental and Heritable Factors in the Causation of Cancer — Analyses of Cohorts of Twins from Sweden, Denmark, and Finland. *New England Journal of Medicine,* 343**,** 78-85.

LIEVRE, A., BACHET, J., BOIGE, V., CAYRE, A., LE CORRE, D., BUC, E., YCHOU, M., BOUCHE, O., LANDI, B. & LOUVET, C. 2008. KRAS mutations as an independent prognostic factor in patients with advanced colorectal cancer treated with cetuximab. *J Clin Oncol,* 26**,** 374 - 379.

LIU, D.-B., GU, Z.-D., CAO, X.-Z., LIU, H. & LI, J.-Y. 2005. Immunocytochemical detection of HoxD9 and Pbx1 homeodomain protein expression in Chinese esophageal squamous cell carcinomas. *World Journal of Gastroenterology,* 11**,** 1562-1566.

LIU, X., JIAN, X. & BOERWINKLE, E. 2011. dbNSFP: A lightweight database of human nonsynonymous SNPs and their functional predictions. *Human Mutation,* 32**,** 894-899.

LIU, Z., ZHU, J., CAO, H., REN, H. & FANG, X. 2012. miR-10b promotes cell invasion through RhoC-AKT signaling pathway by targeting HOXD10 in gastric cancer. *Int J Oncol,* 40**,** 1553-60.

LO, H., WANG, Z., HU, Y., YANG, H., GERE, S., BUETOW, K. & LEE, M. 2003. Allelic variation in gene expression is common in the human genome. *Genome Res,* 13**,** 1855 - 1862.

LONGLEY, M. J., PIERCE, A. J. & MODRICH, P. 1997. DNA polymerase delta is required for human mismatch repair in vitro. *J Biol Chem,* 272**,** 10917-21.

LOPEZ C., T. I. L., KERR P., LORD C.,DEXTER T, IRAVANI M.,ASHWORTH A.,SILVA A. 2006. p53 modulates homologous recombination by transcriptional regulation of the RAD51 gene. *EMBO Rep. ,* 7**,** 219-224.

LOPEZ, R., GARRIDO, E., PINA, P., HIDALGO, A. & LAZOS, M. 2006. HOXB homeobox gene expression in cervical carcinoma. *International Journal of Gynecological Cancer,* 16**,** 329-335.

LORD, C. J. & ASHWORTH, A. 2008. Targeted therapy for cancer using PARP inhibitors. *Curr Opin Pharmacol,* 8**,** 363-9.

LOVEDAY, C., TURNBULL, C., RAMSAY, E., HUGHES, D., RUARK, E., FRANKUM, J. R., BOWDEN, G., KALMYRZAEV, B., WARREN-PERRY, M., SNAPE, K., ADLARD, J. W., BARWELL, J., BERG, J., BRADY, A. F., BREWER, C., BRICE, G., CHAPMAN, C., COOK, J., DAVIDSON, R., DONALDSON, A., DOUGLAS, F., GREENHALGH, L., HENDERSON, A., IZATT, L., KUMAR, A., LALLOO, F., MIEDZYBRODZKA, Z., MORRISON, P. J., PATERSON, J., PORTEOUS, M., ROGERS, M. T., SHANLEY, S., WALKER, L., ECCLES, D., EVANS, D. G., RENWICK, A., SEAL, S., LORD, C. J., ASHWORTH, A., REIS-FILHO, J. S., ANTONIOU, A. C. & RAHMAN, N. 2011. Germline mutations in RAD51D confer susceptibility to ovarian cancer. *Nat Genet,* 43**,** 879-882.

LU, D., KUHN, E., BRISTOW, R. E., GIUNTOLI II, R. L., KJÆR, S. K., SHIH, I.-M. & RODEN, R. B. S. 2011. Comparison of candidate serologic markers for type I and type II ovarian cancer. *Gynecologic Oncology,* 122**,** 560-566.

LU, K. H., WEITZEL, J. N., KODALI, S., WELCH, W. R., BERKOWITZ, R. S. & MOK, S. C. 1997. A novel 4-cM minimally deleted region on chromosome 11p15.1 associated with high grade nonmucinous epithelial ovarian carcinomas. *Cancer Res,* 57**,** 387-90.

LUEDI, P. P., DIETRICH, F. S., WEIDMAN, J. R., BOSKO, J. M., JIRTLE, R. L. & HARTEMINK, A. J. 2007. Computational and experimental identification of novel human imprinted genes. *Genome Research,* 17**,** 1723-1730.

LUO LJ, Z. Z., ZENG JF, LIANG B, YANG JX, CAO DY, SHEN K. 2012. Analysis of the characteristics of side population cells in the human ovarian cancer cell line OVCAR-3. *Zhonghua Fu Chan Ke Za Zhi.,* 47**,** 281-5.

LURIE, G., THOMPSON, P. J., MCDUFFIE, K. E., CARNEY, M. E. & GOODMAN, M. T. 2009. Prediagnostic symptoms of ovarian carcinoma: a case-control study. *Gynecol Oncol,* 114**,** 231-6.

LUTZKER, S. G. & LEVINE, A. J. 1996. A functionally inactive p53 protein in teratocarcinoma cells is activated by either DNA damage or cellular differentiation. *Nat Med,* 2**,** 804-10.

MADORE, J., REN, F., FILALI MOUHIM, A., SANCHEZ, L. & KOEBEL, M. 2009. Characterization of the molecular differences between ovarian endometrioid carcinoma and ovarian serous carcinoma. *The Journal of Pathology,* 220**,** 392-400.

MAEDA, T., TASHIRO, H., KATABUCHI, H., BEGUM, M., OHTAKE, H., KIYONO, T. & OKAMURA, H. 2005. Establishment of an immortalised human ovarian surface epithelial cell line without chromosomal instability. *Br J Cancer,* 93**,** 116-23.

MAGNOWSKA M, S. P., NOWAK-MARKWITZ E, MICHALAK M, MAGNOWSKI P, ROKITA W, KEDZIA H, ZABEL M, SPACZYŃSKI M. 2008. Analysis of hMLH1 and hMSH2 expression in cisplatin-treated ovarian cancer patients. *Ginekol Pol.,* 79**,** 826-34.

MAIA, A.-T., SPITERI, I., LEE, A., O'REILLY, M., JONES, L., CALDAS, C. & PONDER, B. 2009. Extent of differential allelic expression of candidate breast cancer genes is similar in blood and breast. *Breast Cancer Research,* 11**,** R88.

MALLARDO M., P. P., D'URSO O. F. 2008. Non-protein coding RNA biomarkers and differential expression in cancers: a review. *J Exp Clin Cancer Res. ,* 27.

MALUMBRES, M. & BARBACID, M. 2001. To cycle or not to cycle: a critical decision in cancer. *Nat Rev Cancer,* 1**,** 222-31.

MANCEAU, G., KAROUI, M., CHARACHON, A., DELCHIER, J. C. & SOBHANI, I. 2011. [HNPCC (hereditary non-polyposis colorectal cancer) or Lynch syndrome: a syndrome related to a failure of DNA repair system]. *Bull Cancer,* 98**,** 323-36.

MANCINI, D. N., RODENHISER, D. I., AINSWORTH, P. J., O'MALLEY, F. P., SINGH, S. M., XING, W. & ARCHER, T. K. 1998. CpG methylation within the 5' regulatory region of the BRCA1 gene is tumor specific and includes a putative CREB binding site.

Oncogene, 16**,** 1161-9.
MANOHAR, C. F., SALWEN, H. R., FURTADO, M. R. & COHN, S. L. 1996. Up-regulation of HOXC6, HOXD1, and HOXD8 homeobox gene expression in human neuroblastoma cells following chemical induction of differentiation. *Tumour Biol,* 17**,** 34-47.
MARKS, J., DAVIDOFF, A., KERNS, B., HUMPHREY, P., PENCE, J., DODGE, R., CLARKE-PEARSON, D., IGLEHART, J., BAST, R. & BERCHUCK, A. 1991. Overexpression and mutation of p53 in epithelial ovarian cancer. *Cancer Res,* 51**,** 2979 - 2984.
MARTIN, L. P., HAMILTON, T. C. & SCHILDER, R. J. 2008. Platinum Resistance: The Role of DNA Repair Pathways. *Clinical Cancer Research,* 14**,** 1291-1295.
MARTIN, S. A., LORD, C. J. & ASHWORTH, A. 2008. DNA repair deficiency as a therapeutic target in cancer. *Curr Opin Genet Dev,* 18**,** 80-6.
MASSAGUE, J. 2004. G1 cell-cycle control and cancer. *Nature,* 432**,** 298-306.
MASSON, M., NIEDERGANG, C., SCHREIBER, V., MULLER, S., MENISSIER-DE MURCIA, J. & DE MURCIA, G. 1998. XRCC1 is specifically associated with poly(ADP-ribose) polymerase and negatively regulates its activity following DNA damage. *Mol Cell Biol,* 18**,** 3563-71.
MASUTANI, M. & MIWA, M. 2002. Poly(ADP-ribose) Polymerase and Cancer: In Relation to the Lectures Presented by Dr Gilbert de Murcia. *Japanese Journal of Clinical Oncology,* 32**,** 483-487.
MAYER-KUCKUK, P., ULLRICH, O., ZIEGLER, M., GRUNE, T. & SCHWEIGER, M. 1999. Functional interaction of poly(ADP-ribose) with the 20S proteasome in vitro. *Biochem Biophys Res Commun,* 259**,** 576-81.
MAYR, D., HIRSCHMANN, A., LOHRS, U. & DIEBOLD, J. 2006. KRAS and BRAF mutations in ovarian tumors: a comprehensive study of invasive carcinomas, borderline tumors and extraovarian implants. *Gynecol Oncol,* 103**,** 883 - 887.
MAYR, D., KANITZ, V., ANDEREGG, B., LUTHARDT, B., ENGEL, J., LOHRS, U., AMANN, G. & DIEBOLD, J. 2006. Analysis of gene amplification and prognostic markers in ovarian cancer using comparative genomic hybridization for microarrays and immunohistochemical analysis for tissue microarrays. *Am J Clin Pathol,* 126**,** 101-9.
MCCABE, K. M., OLSON, S. B. & MOSES, R. E. 2009. DNA interstrand crosslink repair in mammalian cells. *Journal of cellular physiology,* 220**,** 569-573.
MCINROY, L. & MAATTA, A. 2007. Down-regulation of vimentin expression inhibits carcinoma cell migration and adhesion. *Biochem Biophys Res Commun,* 360**,** 109-14.
MEDEIROS, F., MUTO, M. G., LEE, Y., ELVIN, J. A., CALLAHAN, M. J., FELTMATE, C., GARBER, J. E., CRAMER, D. W. & CRUM, C. P. 2006. The tubal fimbria is a preferred site for early adenocarcinoma in women with familial ovarian cancer syndrome. *Am J Surg Pathol,* 30**,** 230-6.
MEDEIROS, R., PEREIRA, D., AFONSO, N., PALMEIRA, C., FALEIRO, C., AFONSO-LOPES, C., FREITAS-SILVA, M., VASCONCELOS, A., COSTA, S., OSORIO, T. & LOPES, C. 2003. Platinum/paclitaxel-based chemotherapy in advanced ovarian carcinoma: glutathione S-transferase genetic polymorphisms as predictive biomarkers of disease outcome. *Int J Clin Oncol,* 8**,** 156-61.
MELZER, D., PERRY, J. R. B., HERNANDEZ, D., CORSI, A.-M., STEVENS, K., RAFFERTY, I., LAURETANI, F., MURRAY, A., GIBBS, J. R., PAOLISSO, G., RAFIQ, S., SIMON-SANCHEZ, J., LANGO, H., SCHOLZ, S., WEEDON, M. N., AREPALLI, S., RICE, N., WASHECKA, N., HURST, A., BRITTON, A., HENLEY, W., VAN DE LEEMPUT, J., LI, R., NEWMAN, A. B., TRANAH, G., HARRIS, T., PANICKER, V., DAYAN, C., BENNETT, A., MCCARTHY, M. I., RUOKONEN, A., JARVELIN, M.-R., GURALNIK, J., BANDINELLI, S., FRAYLING, T. M., SINGLETON, A. & FERRUCCI, L. 2008. A Genome-Wide Association Study Identifies Protein Quantitative Trait Loci (pQTLs). *PLoS Genet,* 4**,** e1000072.
MEYER, K., MAIA, A., O'REILLY, M., TESCHENDORFF, A., CHIN, S., CALDAS, C. & PONDER, B. 2008. Allele-Specific Up-Regulation of FGFR2 Increases Susceptibility to Breast Cancer. *PLoS Biol,* 6**,** e108.
MEYER, K., MAIA, A.-T., O'REILLY, M., GHOUSSAINI, M. & PRATHALINGAM, R. 2011. A Functional Variant at a Prostate Cancer Predisposition Locus at 8q24 Is Associated with PVT1 Expression. *PLOS Genetics,* 7**,** e1002165.
MIKI, Y., SWENSEN, J., SHATTUCK-EIDENS, D., FUTREAL, P. A., HARSHMAN, K., TAVTIGIAN, S., LIU, Q., COCHRAN, C., BENNETT, L. M., DING, W. & ET AL. 1994. A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. *Science,*

266**,** 66-71.
MILLER, L., SMEDS, J., GEORGE, J., VEGA, V., VERGARA, L., PLONER, A., PAWITAN, Y., HALL, P., KLAAR, S. & LIU, E. 2005. An expression signature for p53 status in human breast cancer predicts mutation status, transcriptional effects, and patient survival. *Proc Natl Acad Sci USA,* 102**,** 13550 - 13555.
MITTAL, K. R., ZELENIUCH-JACQUOTTE, A., COOPER, J. L. & DEMOPOULOS, R. I. 1993. Contralateral ovary in unilateral ovarian carcinoma: a search for preneoplastic lesions. *Int J Gynecol Pathol,* 12**,** 59-63.
MOK, S. C., CHAO, J., SKATES, S., WONG, K.-K., YIU, G. K., MUTO, M. G., BERKOWITZ, R. S. & CRAMER, D. W. 2001. Prostasin, a Potential Serum Marker for Ovarian Cancer: Identification Through Microarray Technology. *Journal of the National Cancer Institute,* 93**,** 1458-1464.
MOLL, R., FRANKE, W. W., SCHILLER, D. L., GEIGER, B. & KREPLER, R. 1982. The catalog of human cytokeratins: patterns of expression in normal epithelia, tumors and cultured cells. *Cell,* 31**,** 11-24.
MOLLOY, P. L. & WATT, F. 1990. DNA methylation and specific protein-DNA interactions. *Philos Trans R Soc Lond B Biol Sci,* 326**,** 267-75.
MONTEIRO, A. N. 2003. BRCA1: the enigma of tissue-specific tumor development. *Trends Genet,* 19**,** 312-5.
MOORE, E. & SOPER, D. E. 1998. Clinical utility of CA125 levels in predicting laparoscopically confirmed salpingitis in patients with clinically diagnosed pelvic inflammatory disease. *Infectious diseases in obstetrics and gynecology,* 6**,** 182-185.
MORGAN, R., PLOWRIGHT, L., HARRINGTON, K., MICHAEL, A. & PANDHA, H. 2010. Targeting HOX and PBX transcription factors in ovarian cancer. *BMC Cancer,* 10**,** 89.
MORLEY, M., MOLONY, C., WEBER, T., DEVLIN, J., EWENS, K., SPIELMAN, R. & CHEUNG, V. 2004. Genetic analysis of genome-wide variation in human gene expression. *Nature,* 430**,** 743 - 747.
MOSMANN, T. 1983. Rapid colorimetric assay for cellular growth and survival: Application to proliferation and cytotoxicity assays. *Journal of Immunological Methods,* 65**,** 55-63.
MOYNAHAN, M. E., CUI, T. Y. & JASIN, M. 2001. Homology-directed DNA repair, mitomycin-C resistance, and chromosome stability is restored with correction of a Brca1 mutation. *Cancer Research,* 61**,** 4842-4850.
MOYNAHAN, M. E., PIERCE, A. J. & JASIN, M. 2001. BRCA2 is required for homology-directed repair of chromosomal breaks. *Molecular cell,* 7**,** 263-272.
MULLENDERS, L. H. & BERNEBURG, M. 2001. Photoimmunology and nucleotide excision repair: impact of transcription coupled and global genome excision repair. *J Photochem Photobiol B,* 65**,** 97-100.
MURDOCH, W. J. & MCDONNEL, A. C. 2002. Roles of the ovarian surface epithelium in ovulation and carcinogenesis. *Reproduction,* 123**,** 743-50.
MUSCHECK, M., SUKOSD, F., PESTI, T. & KOVACS, G. 2000. High density deletion mapping of bladder cancer localizes the putative tumor suppressor gene between loci D8S504 and D8S264 at chromosome 8p23.3. *Lab Invest,* 80**,** 1089-93.
MUSTEA, A., SEHOULI, J., FABJANI, G., KOENSGEN, D., MOBUS, V., BRAICU, E., PIRVULESCU, C., THOMAS, A., TONG, D. & ZEILLINGER, R. 2007. Epidermal growth factor receptor (EGFR) mutation does not correlate with platinum resistance in ovarian carcinoma. Results of a prospective pilot study. *Anticancer Res,* 27**,** 1527 - 1530.
MYERS, C., CHARBONEAU, A., CHEUNG, I., HANKS, D. & BOUDREAU, N. 2002. Sustained expression of homeobox D10 inhibits angiogenesis. *The American journal of pathology,* 161**,** 2099-2109.
NACKLEY, A. G., SHABALINA, S. A., TCHIVILEVA, I. E., SATTERFIELD, K., KORCHYNSKYI, O., MAKAROV, S. S., MAIXNER, W. & DIATCHENKO, L. 2006. Human Catechol-O-Methyltransferase Haplotypes Modulate Protein Expression by Altering mRNA Secondary Structure. *Science,* 314**,** 1930-1933.
NAKAJIMA, T., FUJINO, S., NAKANISHI, G., KIM, Y. S. & JETTEN, A. M. 2004. TIP27: a novel repressor of the nuclear orphan receptor TAK1/TR4. *Nucleic Acids Res,* 32**,** 4194-204.
NAKAMURA, M., YONEKAWA, Y., KLEIHUES, P. & OHGAKI, H. 2001. Promoter hypermethylation of the RB1 gene in glioblastomas. *Lab Invest,* 81**,** 77-82.
NAKAYAMA, N., NAKAYAMA, K., YEASMIN, S., ISHIBASHI, M., KATAGIRI, A., IIDA, K., FUKUMOTO, M. & MIYAZAKI, K. 2008. KRAS or BRAF mutation status is a useful predictor of sensitivity to MEK inhibition in ovarian cancer. *Br J Cancer,* 99**,** 2020 -

2028.
NAM, E. J., YUN, M. J., OH, Y. T., KIM, J. W., KIM, J. H., KIM, S., JUNG, Y. W., KIM, S. W. & KIM, Y. T. 2010. Diagnosis and staging of primary ovarian cancer: correlation between PET/CT, Doppler US, and CT or MRI. *Gynecol Oncol,* 116**,** 389-94.
NAORA, H. 2007. The heterogeneity of epithelial ovarian cancers: reconciling old and new paradigms. *Expert Reviews in Molecular Medicine,* 9**,** 1-12.
NAORA, H., YANG, Y., MONTZ, F. J., SEIDMAN, J. D., KURMAN, R. J. & RODEN, R. B. S. 2001. A serologically identified tumor antigen encoded by a homeobox gene promotes growth of ovarian epithelial cells. *Proceedings of the National Academy of Sciences,* 98**,** 4060-4065.
NAROD, S. A., FEUNTEUN, J., LYNCH, H. T., WATSON, P., CONWAY, T., LYNCH, J. & LENOIR, G. M. 1991. Familial breast-ovarian cancer locus on chromosome 17q12-q23. *Lancet,* 338**,** 82-3.
NAROD, S. A., FORD, D., DEVILEE, P., BARKARDOTTIR, R. B., LYNCH, H. T., SMITH, S. A., PONDER, B. A., WEBER, B. L., GARBER, J. E., BIRCH, J. M. & ET AL. 1995. An evaluation of genetic heterogeneity in 145 breast-ovarian cancer families. Breast Cancer Linkage Consortium. *Am J Hum Genet,* 56**,** 254-64.
NEALE CM, C. R. 1992. Methodology for Genetic Studies of Twins and Families (e-book). *Kluwer Academic Publishers*.
NESS, R. B. & COTTREAU, C. 1999. Possible role of ovarian epithelial inflammation in ovarian cancer. *J Natl Cancer Inst,* 91**,** 1459-67.
NESS, R. B., GRISSO, J. A., VERGONA, R., KLAPPER, J., MORGAN, M. & WHEELER, J. E. 2001. Oral contraceptives, other methods of contraception, and risk reduction for ovarian cancer. *Epidemiology,* 12**,** 307-12.
NETWORK, T. C. G. A. R. 2011. Integrated genomic analyses of ovarian carcinoma. *Nature,* 474**,** 609-615.
NG, G., ROBERTS, I. & COLEMAN, N. 2005. Evaluation of 3 methods of whole-genome amplification for subsequent metaphase comparative genomic hybridization. *Diagn Mol Pathol,* 14**,** 203-12.
NG, P. & HENIKOFF, S. 2006. Predicting the effects of amino acid substitutions on protein function. *Annual review of genomics and human genetics,* 7**,** 61-80.
NIEDERACHER, D., YAN, H. Y., AN, H. X., BENDER, H. G. & BECKMANN, M. W. 1999. CDKN2A gene inactivation in epithelial sporadic ovarian cancer. *Br J Cancer,* 80**,** 1920-6.
NIELSEN, R., PAUL, J. S., ALBRECHTSEN, A. & SONG, Y. S. Genotype and SNP calling from next-generation sequencing data. *Nat Rev Genet,* 12**,** 443-51.
NITTA, R. T., DEL VECCHIO, C. A., CHU, A. H., MITRA, S. S., GODWIN, A. K. & WONG, A. J. 2011. The role of the c-Jun N-terminal kinase 2-[alpha]-isoform in non-small cell lung carcinoma tumorigenesis. *Oncogene,* 30**,** 234-244.
NORD H, H. C., ANDERSSON R, MENZEL U, PFEIFER S, PIOTROWSKI A, BOGDAN A, KLOC W, SANDGREN J, OLOFSSON T, HESSELAGER G, BLOMQUIST E, KOMOROWSKI J, VON DEIMLING A, BRUDER CE, DUMANSKI JP, DÍAZ DE STÅHL T. 2009. Characterization of novel and complex genomic aberrations in glioblastoma using a 32K BAC array. *Neuro Oncol.,* 11**,** 803-18.
NOTARIDOU, M., QUAYE, L., DAFOU, D., JONES, C., SONG, H., HØGDALL, E., KJAER, S. K., CHRISTENSEN, L., HØGDALL, C., BLAAKAER, J., MCGUIRE, V., WU, A. H., VAN DEN BERG, D. J., PIKE, M. C., GENTRY-MAHARAJ, A., WOZNIAK, E., SHER, T., JACOBS, I. J., TYRER, J., SCHILDKRAUT, J. M., MOORMAN, P. G., IVERSEN, E. S., JAKUBOWSKA, A., MĘDREK, K., LUBIŃSKI, J., NESS, R. B., MOYSICH, K. B., LURIE, G., WILKENS, L. R., CARNEY, M. E., WANG-GOHRKE, S., DOHERTY, J. A., ROSSING, M. A., BECKMANN, M. W., THIEL, F. C., EKICI, A. B., CHEN, X., BEESLEY, J., THE AUSTRALIAN OVARIAN CANCER STUDY GROUP/AUSTRALIAN CANCER, S., GRONWALD, J., FASCHING, P. A., CHANG-CLAUDE, J., GOODMAN, M. T., CHENEVIX-TRENCH, G., BERCHUCK, A., PEARCE, C. L., WHITTEMORE, A. S., MENON, U., PHAROAH, P. D. P., GAYTHER, S. A., RAMUS, S. J. & ON BEHALF OF THE OVARIAN CANCER ASSOCIATION, C. 2011. Common alleles in candidate susceptibility genes associated with risk and development of epithelial ovarian cancer. *International Journal of Cancer,* 128**,** 2063-2074.
OBATA, K., MORLAND, S. J., WATSON, R. H., HITCHCOCK, A., CHENEVIX-TRENCH, G., THOMAS, E. J. & CAMPBELL, I. G. 1998. Frequent PTEN/MMAC mutations in

endometrioid but not serous or mucinous epithelial ovarian tumors. *Cancer Res,* 58**,** 2095-7.

OKADA, H., DANOFF, T. M., KALLURI, R. & NEILSON, E. G. 1997. Early role of Fsp1 in epithelial-mesenchymal transformation. *Am J Physiol,* 273**,** F563-74.

OKAMOTO, S., OKAMOTO, A., NIKAIDO, T., SAITO, M., TAKAO, M., YANAIHARA, N., TAKAKURA, S., OCHIAI, K. & TANAKA, T. 2009. Mesenchymal to epithelial transition in the human ovarian surface epithelium focusing on inclusion cysts. *Oncol Rep,* 21**,** 1209-14.

OLSON, S. H., CARLSON, M. D. A., OSTRER, H., HARLAP, S., STONE, A., WINTERS, M. & AMBROSONE, C. B. 2004. Genetic variants in SOD2, MPO, and NQO1, and risk of ovarian cancer. *Gynecologic Oncology,* 93**,** 615-620.

ONO, K., TANAKA, T., TSUNODA, T., KITAHARA, O., KIHARA, C., OKAMOTO, A., OCHIAI, K., TAKAGI, T. & NAKAMURA, Y. 2000. Identification by cDNA Microarray of Genes Involved in Ovarian Carcinogenesis. *Cancer Research,* 60**,** 5007-5011.

OSHER, D. J., DE LEENEER, K., MICHILS, G., HAMEL, N., TOMIAK, E., POPPE, B., LEUNEN, K., LEGIUS, E., SHUEN, A., SMITH, E., ARSENEAU, J., TONIN, P., MATTHIJS, G., CLAES, K., TISCHKOWITZ, M. D. & FOULKES, W. D. 2012. Mutation analysis of RAD51D in non-BRCA1/2 ovarian and breast cancer families. *Br J Cancer,* 106**,** 1460-1463.

OSMAN, N., O'LEARY, N., MULCAHY, E., BARRETT, N., WALLIS, F., HICKEY, K. & GUPTA, R. 2008. Correlation of serum CA125 with stage, grade and survival of patients with epithelial ovarian cancer at a single centre. *Ir Med J,* 101**,** 245-7.

OTSUBO, T., IWAYA, K., MUKAI, Y., MIZOKAMI, Y., SERIZAWA, H., MATSUOKA, T. & MUKAI, K. 2004. Involvement of Arp2/3 complex in the process of colorectal carcinogenesis. *Mod Pathol,* 17**,** 461 - 467.

OZCAN, A., LILES, N., COFFEY, D., SHEN, S. S. & TRUONG, L. D. 2011. PAX2 and PAX8 expression in primary and metastatic mullerian epithelial tumors: a comprehensive comparison. *Am J Surg Pathol,* 35**,** 1837-47.

PAGE, C., LIN, H. J., JIN, Y., CASTLE, V. P., NUNEZ, G., HUANG, M. & LIN, J. 2000. Overexpression of Akt/AKT can modulate chemotherapy-induced apoptosis. *Anticancer Res,* 20**,** 407-16.

PALAYEKAR, M. & HERZOG, T. 2008. The emerging role of epidermal growth factor receptor inhibitors in ovarian cancer. *Int J Gynecol Cancer,* 18**,** 879 - 890.

PANT, P., TAO, H., BEILHARZ, E., BALLINGER, D., COX, D. & FRAZER, K. 2006. Analysis of allelic differential expression in human white blood cells. *Genome Res,* 16**,** 331 - 339.

PARKIN, D. M., BRAY, F., FERLAY, J. & PISANI, P. 2005. Global cancer statistics, 2002. *CA Cancer J Clin,* 55**,** 74-108.

PASTINEN, T., GE, B., GURD, S., GAUDIN, T., DORE, C., LEMIRE, M., LEPAGE, P., HARMSEN, E. & HUDSON, T. 2005. Mapping common regulatory variants to human haplotypes. *Hum Mol Genet,* 14**,** 3963 - 3971.

PASTINEN, T., GE, B. & HUDSON, T. 2006. Influence of human genome polymorphism on gene expression. *Hum Mol Genet,* 15**,** R9 - 16.

PASTINEN, T. & HUDSON, T. 2004. Cis-acting regulatory variation in the human genome. *Science,* 306**,** 647 - 650.

PAVELKA, J. C., LI, A. J. & KARLAN, B. Y. 2007. Hereditary ovarian cancer--assessing risk and prevention strategies. *Obstet Gynecol Clin North Am,* 34**,** 651-65.

PEEDICAYIL, A., VIERKANT, R. A., HARTMANN, L. C., FRIDLEY, B. L., FREDERICKSEN, Z. S., WHITE, K. L., ELLIOTT, E. A., PHELAN, C. M., TSAI, Y.-Y., BERCHUCK, A., IVERSEN, E. S., JR., COUCH, F. J., PEETHAMABARAN, P., LARSON, M. C., KALLI, K. R., KOSEL, M. L., SHRIDHAR, V., RIDER, D. N., LIEBOW, M., CUNNINGHAM, J. M., SCHILDKRAUT, J. M., SELLERS, T. A. & GOODE, E. L. 2010. Risk of Ovarian Cancer and Inherited Variants in Relapse-Associated Genes. *PLoS ONE,* 5**,** e8884.

PENG, H., XU, F., PERSHAD, R., HUNT, K. K., FRAZIER, M. L., BERCHUCK, A., GRAY, J. W., HOGG, D., BAST, R. C., JR. & YU, Y. 2000. ARHI is the center of allelic deletion on chromosome 1p31 in ovarian and breast cancers. *Int J Cancer,* 86**,** 690-4.

PEREGO, P., GIAROLA, M., RIGHETTI, S. C., SUPINO, R., CASERINI, C., DELIA, D., PIEROTTI, M. A., MIYASHITA, T., REED, J. C. & ZUNINO, F. 1996. Association between Cisplatin Resistance and Mutation of p53 Gene and Reduced Bax Expression in Ovarian Carcinoma Cell Systems. *Cancer Research,* 56**,** 556-562.

PEREGO, P., ROMANELLI, S., CARENINI, N., MAGNANI, I., LEONE, R., BONETTI, A.,

PAOLICCHI, A. & ZUNINO, F. 1998. Ovarian cancer cisplatin-resistant cell lines: multiple changes including collateral sensitivity to Taxol. *Ann Oncol,* 9**,** 423-30.

PEREYRA, F., JIA, X., MCLAREN, P. J., TELENTI, A., DE BAKKER, P. I., WALKER, B. D., RIPKE, S., BRUMME, C. J., PULIT, S. L., CARRINGTON, M., KADIE, C. M., CARLSON, J. M., HECKERMAN, D., GRAHAM, R. R., PLENGE, R. M., DEEKS, S. G., GIANNINY, L., CRAWFORD, G., SULLIVAN, J., GONZALEZ, E., DAVIES, L., CAMARGO, A., MOORE, J. M., BEATTIE, N., GUPTA, S., CRENSHAW, A., BURTT, N. P., GUIDUCCI, C., GUPTA, N., GAO, X., QI, Y., YUKI, Y., PIECHOCKA-TROCHA, A., CUTRELL, E., ROSENBERG, R., MOSS, K. L., LEMAY, P., O'LEARY, J., SCHAEFER, T., VERMA, P., TOTH, I., BLOCK, B., BAKER, B., ROTHCHILD, A., LIAN, J., PROUDFOOT, J., ALVINO, D. M., VINE, S., ADDO, M. M., ALLEN, T. M., ALTFELD, M., HENN, M. R., LE GALL, S., STREECK, H., HAAS, D. W., KURITZKES, D. R., ROBBINS, G. K., SHAFER, R. W., GULICK, R. M., SHIKUMA, C. M., HAUBRICH, R., RIDDLER, S., SAX, P. E., DAAR, E. S., RIBAUDO, H. J., AGAN, B., AGARWAL, S., AHERN, R. L., ALLEN, B. L., ALTIDOR, S., ALTSCHULER, E. L., AMBARDAR, S., ANASTOS, K., ANDERSON, B., ANDERSON, V., ANDRADY, U., ANTONISKIS, D., BANGSBERG, D., BARBARO, D., BARRIE, W., BARTCZAK, J., BARTON, S., BASDEN, P., BASGOZ, N., BAZNER, S., BELLOS, N. C., BENSON, A. M., BERGER, J., BERNARD, N. F., BERNARD, A. M., BIRCH, C., BODNER, S. J., BOLAN, R. K., BOUDREAUX, E. T., BRADLEY, M., BRAUN, J. F., BRNDJAR, J. E., BROWN, S. J., BROWN, K., BROWN, S. T., et al. 2010. The major genetic determinants of HIV-1 control affect HLA class I peptide presentation. *Science,* 330**,** 1551-7.

PERL, A.-K., WILGENBUS, P., DAHL, U., SEMB, H. & CHRISTOFORI, G. 1998. A causal role for E-cadherin in the transition from adenoma to carcinoma. *Nature,* 392**,** 190-193.

PFEIDERER, A. 1984. [Biology of ovarian carcinoma and the hope of individualised therapy]. *Onkologie,* 2**,** 82-8.

PHAROAH, P., TYRER, J., DUNNING, A., EASTON, D., PONDER, B. & INVESTIGATORS, S. 2007. Team RDC: R: A Language and Environment for Statistical Computing

Association between common variation in 120 candidate genes and breast cancer risk. *PLoS Genet,* 3**,** e42.

PHAROAH, P. D., EASTON, D. F., STOCKTON, D. L., GAYTHER, S. & PONDER, B. A. 1999. Survival in familial, BRCA1-associated, and BRCA2-associated epithelial ovarian cancer. United Kingdom Coordinating Committee for Cancer Research (UKCCCR) Familial Ovarian Cancer Study Group. *Cancer Res,* 59**,** 868-71.

PHAROAH, P. D. & PONDER, B. A. 2002. The genetics of ovarian cancer. *Best Pract Res Clin Obstet Gynaecol,* 16**,** 449-68.

PHILLIPS, N. J., ZIEGLER, M. R., RADFORD, D. M., FAIR, K. L., STEINBRUECK, T., XYNOS, F. P. & DONIS-KELLER, H. 1996. Allelic deletion on chromosome 17p13.3 in early ovarian cancer. *Cancer Res,* 56**,** 606-11.

PIEK, J., TORRENGA, B., HERMSEN, B., VERHEIJEN, R., ZWEEMER, R., GILLE, J., KENEMANS, P., VAN DIEST, P. & MENKO, F. 2003. Histopathological characteristics of <i>BRCA1-</i> and <i>BRCA2</i>-associated intraperitoneal cancer: a clinic-based study. *Familial Cancer,* 2**,** 73-78.

PIEK, J., VAN DIEST, P., ZWEEMER, R., JANSEN, J., POORT-KEESOM, R., MENKO, F., GILLE, J., JONGSMA, A., PALS, G. & KENEMANS, P. 2001. Dysplastic changes in prophylactically removed Fallopian tubes of women predisposed to developing ovarian cancer. *J Pathol,* 195**,** 451 - 456.

PIEK, J. M., KENEMANS, P. & VERHEIJEN, R. H. 2004. Intraperitoneal serous adenocarcinoma: a critical appraisal of three hypotheses on its cause. *Am J Obstet Gynecol,* 191**,** 718-32.

PIERETTI, M., HOPENHAYN-RICH, C., KHATTAR, N. H., CAO, Y., HUANG, B. & TUCKER, T. C. 2002. Heterogeneity of ovarian cancer: relationships among histological group, stage of disease, tumor markers, patient characteristics, and survival. *Cancer Invest,* 20**,** 11-23.

PLEASANCE, E. D., STEPHENS, P. J., O'MEARA, S., MCBRIDE, D. J., MEYNERT, A., JONES, D., LIN, M.-L., BEARE, D., LAU, K. W., GREENMAN, C., VARELA, I., NIK-ZAINAL, S., DAVIES, H. R., ORDONEZ, G. R., MUDIE, L. J., LATIMER, C., EDKINS, S., STEBBINGS, L., CHEN, L., JIA, M., LEROY, C., MARSHALL, J., MENZIES, A., BUTLER, A., TEAGUE, J. W., MANGION, J., SUN, Y. A., MCLAUGHLIN, S. F., PECKHAM, H. E., TSUNG, E. F., COSTA, G. L., LEE, C. C., MINNA, J. D., GAZDAR,

A., BIRNEY, E., RHODES, M. D., MCKERNAN, K. J., STRATTON, M. R., FUTREAL, P. A. & CAMPBELL, P. J. 2010. A small-cell lung cancer genome with complex signatures of tobacco exposure. *Nature,* 463**,** 184-190.

PLESCHKE, J. M., KLECZKOWSKA, H. E., STROHM, M. & ALTHAUS, F. R. 2000. Poly(ADP-ribose) binds to specific domains in DNA damage checkpoint proteins. *J Biol Chem,* 275**,** 40974-80.

PLONER, A., MILLER, L., HALL, P., BERGH, J. & PAWITAN, Y. 2005. Correlation test to assess low-level processing of high-density oligonucleotide microarray data. *BMC Bioinformatics,* 6**,** 80.

PLUMB, J. A., STRATHDEE, G., SLUDDEN, J., KAYE, S. B. & BROWN, R. 2000. Reversal of drug resistance in human tumor xenografts by 2'-deoxy-5-azacytidine-induced demethylation of the hMLH1 gene promoter. *Cancer Res,* 60**,** 6039-44.

POHL, G., HO, C.-L., KURMAN, R. J., BRISTOW, R., WANG, T.-L. & SHIH, I.-M. 2005. Inactivation of the Mitogen-Activated Protein Kinase Pathway as a Potential Target-Based Therapy in Ovarian Serous Tumors with KRAS or BRAF Mutations. *Cancer Research,* 65**,** 1994-2000.

POLAKIS, P. 2000. Wnt signaling and cancer. *Genes & Development,* 14**,** 1837-1851.

POON, S. L., KLAUSEN, C., HAMMOND, G. L. & LEUNG, P. C. 2011. 37-kDa laminin receptor precursor mediates GnRH-II-induced MMP-2 expression and invasiveness in ovarian cancer cells. *Mol Endocrinol,* 25**,** 327-38.

POTHURI, B., LEITAO, M. M., LEVINE, D. A., VIALE, A., OLSHEN, A. B., ARROYO, C., BOGOMOLNIY, F., OLVERA, N., LIN, O., SOSLOW, R. A., ROBSON, M. E., OFFIT, K., BARAKAT, R. R. & BOYD, J. 2010. Genetic Analysis of the Early Natural History of Epithelial Ovarian Carcinoma. *PLoS ONE,* 5**,** e10358.

POWELL, C. B., KENLEY, E., CHEN, L. M., CRAWFORD, B., MCLENNAN, J., ZALOUDEK, C., KOMAROMY, M., BEATTIE, M. & ZIEGLER, J. 2005. Risk-reducing salpingo-oophorectomy in BRCA mutation carriers: role of serial sectioning in the detection of occult malignancy. *J Clin Oncol,* 23**,** 127-32.

PRESS, M. F., NOUSEK-GOEBL, N. A., BUR, M. & GREENE, G. L. 1986. Estrogen receptor localization in the female genital tract. *Am J Pathol,* 123**,** 280-92.

PRICE, A. L., HELGASON, A., THORLEIFSSON, G., MCCARROLL, S. A., KONG, A. & STEFANSSON, K. 2011. Single-tissue and cross-tissue heritability of gene expression via identity-by-descent in related or unrelated individuals. *PLoS Genet,* 7**,** 24.

PRICE, A. L., PATTERSON, N., HANCKS, D. C., MYERS, S., REICH, D., CHEUNG, V. G. & SPIELMAN, R. S. 2008. Effects of cis and trans Genetic Ancestry on Gene Expression in African Americans. *PLoS Genet,* 4**,** e1000294.

PRIX, L., UCIECHOWSKI, P., BOCKMANN, B., GIESING, M. & SCHUETZ, A. 2002. Diagnostic biochip array for fast and sensitive detection of K-ras mutations in stool. *Clin Chem,* 48**,** 428 - 435.

PROIA, T. A., KELLER, P. J., GUPTA, P. B., KLEBBA, I., JONES, A. D., SEDIC, M., GILMORE, H., TUNG, N., NABER, S. P., SCHNITT, S., LANDER, E. S. & KUPERWASSER, C. 2010. Genetic predisposition directs breast cancer phenotype by dictating progenitor cell fate. *Cell Stem Cell,* 8**,** 149-63.

QUINN, M. C. J., FILALI MOUHIM, A., PROVENCHER, D., MES MASSON, A.-M. & TONIN, P. 2009. Reprogramming of the Transcriptome in a Novel Chromosome 3 Transfer Tumor Suppressor Ovarian Cancer Cell Line Model Affected Molecular Networks That Are Characteristic of Ovarian Cancer. *Molecular carcinogenesis,* 48**,** 648-661.

RACILA, E., RACILA, D., RITCHIE, J., TAYLOR, C., DAHLE, C. & WEINER, G. 2006. The pattern of clinical breast cancer metastasis correlates with a single nucleotide polymorphism in the C1qA component of complement. *Immunogenetics,* 58**,** 1 - 8.

RAMUS, S. J., ANTONIOU, A. C., KUCHENBAECKER, K. B., SOUCY, P., BEESLEY, J., CHEN, X., MCGUFFOG, L., SINILNIKOVA, O. M., HEALEY, S., BARROWDALE, D., LEE, A., THOMASSEN, M., GERDES, A. M., KRUSE, T. A., JENSEN, U. B., SKYTTE, A. B., CALIGO, M. A., LILJEGREN, A., LINDBLOM, A., OLSSON, H., KRISTOFFERSSON, U., STENMARK-ASKMALM, M., MELIN, B., DOMCHEK, S. M., NATHANSON, K. L., REBBECK, T. R., JAKUBOWSKA, A., LUBINSKI, J., JAWORSKA, K., DURDA, K., ZLOWOCKA, E., GRONWALD, J., HUZARSKI, T., BYRSKI, T., CYBULSKI, C., TOLOCZKO-GRABAREK, A., OSORIO, A., BENITEZ, J., DURAN, M., TEJADA, M. I., HAMANN, U., ROOKUS, M., VAN LEEUWEN, F. E., AALFS, C. M., MEIJERS-HEIJBOER, H. E., VAN ASPEREN, C. J., VAN ROOZENDAAL, K. E.,

HOOGERBRUGGE, N., COLLEE, J. M., KRIEGE, M., VAN DER LUIJT, R. B., PEOCK, S., FROST, D., ELLIS, S. D., PLATTE, R., FINEBERG, E., EVANS, D. G., LALLOO, F., JACOBS, C., EELES, R., ADLARD, J., DAVIDSON, R., ECCLES, D., COLE, T., COOK, J., PATERSON, J., DOUGLAS, F., BREWER, C., HODGSON, S., MORRISON, P. J., WALKER, L., PORTEOUS, M. E., KENNEDY, M. J., PATHAK, H., GODWIN, A. K., STOPPA-LYONNET, D., CAUX-MONCOUTIER, V., DE PAUW, A., GAUTHIER-VILLARS, M., MAZOYER, S., LEONE, M., CALENDER, A., LASSET, C., BONADONA, V., HARDOUIN, A., BERTHET, P., BIGNON, Y. J., UHRHAMMER, N., FAIVRE, L., LOUSTALOT, C., BUYS, S., DALY, M., MIRON, A., TERRY, M. B., CHUNG, W. K., JOHN, E. M., SOUTHEY, M., GOLDGAR, D., SINGER, C. F., TEA, M. K., et al. 2012. Ovarian cancer susceptibility alleles and risk of ovarian cancer in BRCA1 and BRCA2 mutation carriers. *Hum Mutat,* 33**,** 690-702.

RAMUS, S. J., FISHMAN, A., PHAROAH, P. D., YARKONI, S., ALTARAS, M. & PONDER, B. A. 2001. Ovarian cancer survival in Ashkenazi Jewish patients with BRCA1 and BRCA2 mutations. *Eur J Surg Oncol,* 27**,** 278-81.

RAMUS, S. J. & GAYTHER, S. A. 2009. The contribution of BRCA1 and BRCA2 to ovarian cancer. *Mol Oncol,* 3**,** 138-50.

RAMUS, S. J., HARRINGTON, P. A., PYE, C., DICIOCCIO, R. A., COX, M. J., GARLINGHOUSE-JONES, K., OAKLEY-GIRVAN, I., JACOBS, I. J., HARDY, R. M., WHITTEMORE, A. S., PONDER, B. A., PIVER, M. S., PHAROAH, P. D. & GAYTHER, S. A. 2007. Contribution of BRCA1 and BRCA2 mutations to inherited ovarian cancer. *Hum Mutat,* 28**,** 1207-15.

RAMUS, S. J., PHAROAH, P. D., HARRINGTON, P., PYE, C., WERNESS, B., BOBROW, L., AYHAN, A., WELLS, D., FISHMAN, A., GORE, M., DICIOCCIO, R. A., PIVER, M. S., WHITTEMORE, A. S., PONDER, B. A. & GAYTHER, S. A. 2003. BRCA1/2 mutation status influences somatic genetic progression in inherited and sporadic epithelial ovarian cancer cases. *Cancer Res,* 63**,** 417-23.

RAUCH, T., ZHONG, X., WU, X., WANG, M. & KERNSTINE, K. 2008. High-resolution mapping of DNA hypermethylation and hypomethylation in lung cancer. *Proceedings of the National Academy of Sciences of the United States of America,* 105**,** 252-257.

RAYCHAUDHURI, S. 2011. Mapping rare and common causal alleles for complex human diseases. *Cell,* 147**,** 57-69.

REBBECK, T. R., MITRA, N., DOMCHEK, S. M., WAN, F., CHUAI, S., FRIEBEL, T. M., PANOSSIAN, S., SPURDLE, A., CHENEVIX-TRENCH, G., SINGER, C. F., PFEILER, G., NEUHAUSEN, S. L., LYNCH, H. T., GARBER, J. E., WEITZEL, J. N., ISAACS, C., COUCH, F., NAROD, S. A., RUBINSTEIN, W. S., TOMLINSON, G. E., GANZ, P. A., OLOPADE, O. I., TUNG, N., BLUM, J. L., GREENBERG, R., NATHANSON, K. L. & DALY, M. B. 2009. Modification of ovarian cancer risk by BRCA1/2-interacting genes in a multicenter cohort of BRCA1/2 mutation carriers. *Cancer Res,* 69**,** 5801-10.

REBBECK, T. R., MITRA, N., DOMCHEK, S. M., WAN, F., FRIEBEL, T. M., TRAN, T. V., SINGER, C. F., TEA, M.-K. M., BLUM, J. L., TUNG, N., OLOPADE, O. I., WEITZEL, J. N., LYNCH, H. T., SNYDER, C. L., GARBER, J. E., ANTONIOU, A. C., PEOCK, S., EVANS, D. G., PATERSON, J., KENNEDY, M. J., DONALDSON, A., DORKINS, H., EASTON, D. F., RUBINSTEIN, W. S., DALY, M. B., ISAACS, C., NEVANLINNA, H., COUCH, F. J., ANDRULIS, I. L., FREIDMAN, E., LAITMAN, Y., GANZ, P. A., TOMLINSON, G. E., NEUHAUSEN, S. L., NAROD, S. A., PHELAN, C. M., GREENBERG, R. & NATHANSON, K. L. 2011. Modification of BRCA1-Associated Breast and Ovarian Cancer Risk by BRCA1-Interacting Genes. *Cancer Research,* 71**,** 5792-5805.

REIK, W., DEAN, W. & WALTER, J. 2001. Epigenetic reprogramming in mammalian development. *Science,* 293**,** 1089-93.

RELES, A., WEN, W., SCHMIDER, A., GEE, C., RUNNEBAUM, I., KILIAN, U., JONES, L., EL-NAGGAR, A., MINGUILLON, C. & SCHONBORN, I. 2001. Correlation of p53 mutations with resistance to platinum-based chemotherapy and shortened survival in ovarian cancer. *Clin Cancer Res,* 7**,** 2984 - 2997.

REMY, I. & MICHNICK, S. W. 2004. Regulation of apoptosis by the Ft1 protein, a new modulator of protein kinase B/Akt. *Mol Cell Biol,* 24**,** 1493-504.

REUMERS, J., CONDE, L., MEDINA, I., MAURER-STROH, S., VAN DURME, J., DOPAZO, J., ROUSSEAU, F. & SCHYMKOWITZ, J. 2008. Joint annotation of coding and non-coding single nucleotide polymorphisms and mutations in the SNPeffect and PupaSuite

databases. *Nucleic Acids Res,* 36**,** 17.
RISCH, H. A. 1998. Hormonal Etiology of Epithelial Ovarian Cancer, With a Hypothesis Concerning the Role of Androgens and Progesterone. *Journal of the National Cancer Institute,* 90**,** 1774-1786.
RISCH, H. A., MARRETT, L. D. & HOWE, G. R. 1994. Parity, contraception, infertility, and the risk of epithelial ovarian cancer. *Am J Epidemiol,* 140**,** 585-97.
RISCH, H. A., MCLAUGHLIN, J. R., COLE, D. E., ROSEN, B., BRADLEY, L., FAN, I., TANG, J., LI, S., ZHANG, S., SHAW, P. A. & NAROD, S. A. 2006. Population BRCA1 and BRCA2 mutation frequencies and cancer penetrances: a kin-cohort study in Ontario, Canada. *J Natl Cancer Inst,* 98**,** 1694-706.
RISCH, N. & MERIKANGAS, K. 1996. The Future of Genetic Studies of Complex Human Diseases. *Science,* 273**,** 1516-1517.
ROH, M. H., YASSIN, Y., MIRON, A., MEHRA, K. K., MEHRAD, M., MONTE, N. M., MUTTER, G. L., NUCCI, M. R., NING, G., MCKEON, F. D., HIRSCH, M. S., WA, X. & CRUM, C. P. 2010. High-grade fimbrial-ovarian carcinomas are unified by altered p53, PTEN and PAX2 expression. *Mod Pathol*.
ROSENBLATT, K. A., WEISS, N. S., CUSHING-HAUGEN, K. L., WICKLUND, K. G. & ROSSING, M. A. 2011. Genital powder exposure and the risk of epithelial ovarian cancer. *Cancer Causes Control,* 22**,** 737-42.
ROSSING, M. A., DALING, J. R., WEISS, N. S., MOORE, D. E. & SELF, S. G. 1994. Ovarian Tumors in a Cohort of Infertile Women. *New England Journal of Medicine,* 331**,** 771-776.
ROY, R., CHUN, J. & POWELL, S. N. 2012. BRCA1 and BRCA2: different roles in a common pathway of genome protection. *Nat Rev Cancer,* 12**,** 68-78.
RUBIN, S. C., BENJAMIN, I., BEHBAKHT, K., TAKAHASHI, H., MORGAN, M. A., LIVOLSI, V. A., BERCHUCK, A., MUTO, M. G., GARBER, J. E., WEBER, B. L., LYNCH, H. T. & BOYD, J. 1996. Clinical and pathological features of ovarian cancer in women with germ-line mutations of BRCA1. *N Engl J Med,* 335**,** 1413-6.
RUSSELL, S. & MCCLUGGAGE, W. 2004. A multistep model for ovarian tumorigenesis: the value of mutation analysis in the KRAS and BRAF genes. *J Pathol,* 203**,** 617 - 619.
RYAN, J., KAUFMANN, W., RAAB-TRAUB, N., OGLESBEE, S., CAREY, L. & GULLEY, M. 2006. Clonal evolution of lymphoblastoid cell lines. *Lab Invest,* 86**,** 1193 - 1200.
SAGAE, S., KOBAYASHI, K., NISHIOKA, Y., SUGIMURA, M., ISHIOKA, S., NAGATA, M., TERASAWA, K., TOKINO, T. & KUDO, R. 1999. Mutational analysis of beta-catenin gene in Japanese ovarian carcinomas: frequent mutations in endometrioid carcinomas. *Jpn J Cancer Res,* 90**,** 510-5.
SAITO, M., OKAMOTO, A., KOHNO, T., TAKAKURA, S., SHINOZAKI, H., ISONISHI, S., YASUHARA, T., YOSHIMURA, T., OHTAKE, Y., OCHIAI, K., YOKOTA, J. & TANAKA, T. 2000. Allelic imbalance and mutations of the PTEN gene in ovarian cancer. *Int J Cancer,* 85**,** 160-5.
SAJJAD, Y. 2010. Development of the genital ducts and external genitalia in the early human embryo. *J Obstet Gynaecol Res,* 36**,** 929-37.
SALAHSHOR, S. & WOODGETT, J. R. 2005. The links between axin and carcinogenesis. *Journal of Clinical Pathology,* 58**,** 225-236.
SALAMANCA, C. M., MAINES-BANDIERA, S. L., LEUNG, P. C., HU, Y. L. & AUERSPERG, N. 2004. Effects of epidermal growth factor/hydrocortisone on the growth and differentiation of human ovarian surface epithelium. *J Soc Gynecol Investig,* 11**,** 241-51.
SATO, N., TSUNODA, H., NISHIDA, M., MORISHITA, Y., TAKIMOTO, Y., KUBO, T. & NOGUCHI, M. 2000. Loss of heterozygosity on 10q23.3 and mutation of the tumor suppressor gene PTEN in benign endometrial cyst of the ovary: possible sequence progression from benign endometrial cyst to endometrioid carcinoma and clear cell carcinoma of the ovary. *Cancer Res,* 60**,** 7052-6.
SATO, T., SAITO, H., MORITA, R., KOI, S., LEE, J. H. & NAKAMURA, Y. 1991. Allelotype of human ovarian cancer. *Cancer Res,* 51**,** 5118-22.
SCHADT, E., MONKS, S., DRAKE, T., LUSIS, A., CHE, N., COLINAYO, V., RUFF, T., MILLIGAN, S., LAMB, J., CAVET, G., LINSLEY, P., MAO, M., STOUGHTON, R. & FRIEND, S. 2003. Genetics of gene expression surveyed in maize, mouse and man. *Nature,* 422**,** 297 - 302.
SCHARER, O. D. & JIRICNY, J. 2001. Recent progress in the biology, chemistry and structural biology of DNA glycosylases. *Bioessays,* 23**,** 270-81.

SCHARL, A., CROMBACH, G., VIERBUCHEN, M., MÜSCH, H. & BOLTE, A. 1989. CA 125 in normal tissues and carcinomas of the uterine cervix, endometrium and fallopian tube. *Archives of Gynecology and Obstetrics,* 244**,** 103-112.

SCHILDER, R. J., SILL, M. W., CHEN, X., DARCY, K. M., DECESARE, S. L., LEWANDOWSKI, G., LEE, R. B., ARCIERO, C. A., WU, H. & GODWIN, A. K. 2005. Phase II study of gefitinib in patients with relapsed or persistent ovarian or primary peritoneal carcinoma and evaluation of epidermal growth factor receptor mutations and immunohistochemical expression: a Gynecologic Oncology Group Study. *Clin Cancer Res,* 11**,** 5539-48.

SCHILDKRAUT, J. M., GOODE, E. L., CLYDE, M. A., IVERSEN, E. S., MOORMAN, P. G., BERCHUCK, A., MARKS, J. R., LISSOWSKA, J., BRINTON, L., PEPLONSKA, B., CUNNINGHAM, J. M., VIERKANT, R. A., RIDER, D. N., CHENEVIX-TRENCH, G., WEBB, P. M., BEESLEY, J., CHEN, X., PHELAN, C., SUTPHEN, R., SELLERS, T. A., PEARCE, L., WU, A. H., VAN DEN BERG, D., CONTI, D., ELUND, C. K., ANDERSON, R., GOODMAN, M. T., LURIE, G., CARNEY, M. E., THOMPSON, P. J., GAYTHER, S. A., RAMUS, S. J., JACOBS, I., KRUGER KJAER, S., HOGDALL, E., BLAAKAER, J., HOGDALL, C., EASTON, D. F., SONG, H., PHAROAH, P. D., WHITTEMORE, A. S., MCGUIRE, V., QUAYE, L., ANTON-CULVER, H., ZIOGAS, A., TERRY, K. L., CRAMER, D. W., HANKINSON, S. E., TWOROGER, S. S., CALINGAERT, B., CHANOCK, S., SHERMAN, M. & GARCIA-CLOSAS, M. 2009. Single nucleotide polymorphisms in the TP53 region and susceptibility to invasive epithelial ovarian cancer. *Cancer Res,* 69**,** 2349-57.

SCHREIBER, V., AME, J. C., DOLLE, P., SCHULTZ, I., RINALDI, B., FRAULOB, V., MENISSIER-DE MURCIA, J. & DE MURCIA, G. 2002. Poly(ADP-ribose) polymerase-2 (PARP-2) is required for efficient base excision DNA repair in association with PARP-1 and XRCC1. *J Biol Chem,* 277**,** 23028-36.

SCHROCK, E., DU MANOIR, S., VELDMAN, T., SCHOELL, B., WIENBERG, J., FERGUSON-SMITH, M. A., NING, Y., LEDBETTER, D. H., BAR-AM, I., SOENKSEN, D., GARINI, Y. & RIED, T. 1996. Multicolor spectral karyotyping of human chromosomes. *Science,* 273**,** 494-7.

SCHULTZ, R. A., SAXON, P. J., GLOVER, T. W. & FRIEDBERG, E. C. 1987. Microcell-mediated transfer of a single human chromosome complements xeroderma pigmentosum group A fibroblasts. *Proceedings of the National Academy of Sciences,* 84**,** 4176-4179.

SCHUYER, M., HENZEN-LOGMANS, S., BURG, M., FIERET, J., DERKSEN, C., LOOK, M., MEIJER-VAN GELDER, M., KLIJN, J., FOEKENS, J. & BERNS, E. 1999. Genetic alterations in ovarian borderline tumours and ovarian carcinomas. *Eur J Obstet Gynecol Reprod Biol,* 82**,** 147 - 150.

SCHWARTZ, S., JR., YAMAMOTO, H., NAVARRO, M., MAESTRO, M., REVENTOS, J. & PERUCHO, M. 1999. Frameshift mutations at mononucleotide repeats in caspase-5 and other target genes in endometrial and gastrointestinal cancer of the microsatellite mutator phenotype. *Cancer Res,* 59**,** 2995-3002.

SCIAN, M., STAGLIANO, K., DEB, D., ELLIS, M., CARCHMAN, E., DAS, A., VALERIE, K., DEB, S. & DEB, S. 2004. Tumor-derived p53 mutants induce oncogenesis by transactivating growth-promoting genes. *Oncogene,* 23**,** 4430 - 4443.

SCULLY, R., CHEN, J., PLUG, A., XIAO, Y., WEAVER, D., FEUNTEUN, J., ASHLEY, T. & LIVINGSTON, D. M. 1997. Association of BRCA1 with Rad51 in mitotic and meiotic cells. *Cell,* 88**,** 265-75.

SERRE, D., GURD, S., GE, B., SLADEK, R., SINNETT, D., HARMSEN, E., BIBIKOVA, M., CHUDIN, E., BARKER, D., DICKINSON, T., FAN, J. & HUDSON, T. 2008. Differential allelic expression in the human genome: a robust approach to identify genetic and epigenetic cis-acting mechanisms regulating gene expression. *PLoS Genet,* 4**,** e1000006.

SEVIN, B. U. & PERRAS, J. P. 1997. Tumor heterogeneity and in vitro chemosensitivity testing in ovarian cancer. *Am J Obstet Gynecol,* 176**,** 759-66.

SHAH, P. P. & KAKAR, S. S. Pituitary tumor transforming gene induces epithelial to mesenchymal transition by regulation of Twist, Snail, Slug, and E-cadherin. *Cancer Lett,* 311**,** 66-76.

SHAH, P. P. & KAKAR, S. S. 2011. Pituitary tumor transforming gene induces epithelial to mesenchymal transition by regulation of Twist, Snail, Slug, and E-cadherin. *Cancer Letters,* 311**,** 66-76.

SHAHIN, M., HUGHES, J., SOOD, A. & BULLER, R. 2000. The prognostic significance of p53 tumor suppressor gene alterations in ovarian carcinoma. *Cancer,* 89**,** 2006 - 2017.

SHAN, W. & LIU, J. 2009. Epithelial ovarian cancer: focus on genetics and animal models. *Cell Cycle,* 8**,** 731-5.

SHAN W., M.-U. I., ZHANG J., ROSEN D., ZHANG S., WEI J. AND LIU J. 2012. Mucinous adenocarcinoma developed from human fallopian tube epithelial cells through defined genetic modifications. *Cell cycle,* 11**,** 2107 - 2113.

SHAO, G., PATTERSON-FORTIN, J., MESSICK, T. E., FENG, D., SHANBHAG, N., WANG, Y. & GREENBERG, R. A. 2009. MERIT40 controls BRCA1-Rap80 complex integrity and recruitment to DNA double-strand breaks. *Genes Dev,* 23**,** 740-54.

SHAYESTEH, L., LU, Y., KUO, W. L., BALDOCCHI, R., GODFREY, T., COLLINS, C., PINKEL, D., POWELL, B., MILLS, G. B. & GRAY, J. W. 1999. PIK3CA is implicated as an oncogene in ovarian cancer. *Nat Genet,* 21**,** 99-102.

SHENSON, D. L., GALLION, H. H., POWELL, D. E. & PIERETTI, M. 1995. Loss of heterozygosity and genomic instability in synchronous endometrioid tumors of the ovary and endometrium. *Cancer,* 76**,** 650-7.

SHIH, Y. C., KERR, J., LIU, J., HURST, T., KHOO, S. K., WARD, B., WAINWRIGHT, B. & CHENEVIX-TRENCH, G. 1997. Rare mutations and no hypermethylation at the CDKN2A locus in epithelial ovarian tumours. *Int J Cancer,* 70**,** 508-11.

SHIMIZU, Y. 1997. Individualized chemotherapeutic regimen for each histological subtype of ovarian carcinoma. *Gan To Kagaku Ryoho,* 1**,** 61-9.

SHIVJI, M. K. K., MUKUND, S. R., RAJENDRA, E., CHEN, S., SHORT, J. M., SAVILL, J., KLENERMAN, D. & VENKITARAMAN, A. R. 2009. The BRC repeats of human BRCA2 differentially regulate RAD51 binding on single- versus double-stranded DNA to stimulate strand exchange. *Proceedings of the National Academy of Sciences,* 106**,** 13254-13259.

SIEBEN, N., MACROPOULOS, P., ROEMEN, G., KOLKMAN-ULJEE, S., JAN FLEUREN, G., HOUMADI, R., DISS, T., WARREN, B., AL ADNANI, M. & DE GOEIJ, A. 2004. In ovarian neoplasms, BRAF, but not KRAS, mutations are restricted to low-grade serous tumours. *J Pathol,* 202**,** 336 - 340.

SIMON A. GAYTHER, P. R., PATRICIA HARRINGTON, ANTONIS C. ANTONIOU, & DOUGLAS F. EASTON, A. B. A. J. P. 1999. *The Contribution of Germline BRCA1 and BRCA2 Mutations to Familial*

Ovarian Cancer: No Evidence for Other Ovarian Cancer–Susceptibility

Genes [Online]. 65].

SINGER, G., STOHR, R., COPE, L., DEHARI, R., HARTMANN, A., CAO, D. F., WANG, T. L., KURMAN, R. J. & SHIH IE, M. 2005. Patterns of p53 mutations separate ovarian serous borderline tumors and low- and high-grade carcinomas and provide support for a new model of ovarian carcinogenesis: a mutational analysis with immunohistochemical correlation. *Am J Surg Pathol,* 29**,** 218-24.

SIRBU, B. M., LACHMAYER, S. J., WÜLFING, V., MARTEN, L. M., CLARKSON, K. E., LEE, L. W., GHEORGHIU, L., ZOU, L., POWELL, S. N., DAHM-DAPHI, J. & WILLERS, H. 2011. ATR-p53 Restricts Homologous Recombination in Response to Replicative Stress but Does Not Limit DNA Interstrand Crosslink Repair in Lung Cancer Cells. *PLoS ONE,* 6**,** e23053.

SKIRNISDOTTIR, I., SORBE, B. & SEIDAL, T. 2001. The growth factor receptors HER-2/neu and EGFR, their relationship, and their effects on the prognosis in early stage (FIGO I-II) epithelial ovarian carcinoma. *Int J Gynecol Cancer,* 11**,** 119-29.

SMALL, K. S., HEDMAN, A. K., GRUNDBERG, E., NICA, A. C., THORLEIFSSON, G., KONG, A., THORSTEINDOTTIR, U., SHIN, S. Y., RICHARDS, H. B., SORANZO, N., AHMADI, K. R., LINDGREN, C. M., STEFANSSON, K., DERMITZAKIS, E. T., DELOUKAS, P., SPECTOR, T. D. & MCCARTHY, M. I. 2011. Identification of an imprinted master trans regulator at the KLF14 locus related to multiple metabolic phenotypes. *Nat Genet,* 43**,** 561-4.

SONG, H., RAMUS, S., QUAYE, L., DICIOCCIO, R., TYRER, J., LOMAS, E., SHADFORTH, D., HOGDALL, E., HOGDALL, C., MCGUIRE, V., WHITTEMORE, A., EASTON, D., PONDER, B., KJAER, S., PHAROAH, P. & GAYTHER, S. 2006. Common variants in mismatch repair genes and risk of invasive ovarian cancer. *Carcinogenesis,* 27**,** 2235 - 2242.

SONG, H., RAMUS, S. J., QUAYE, L., DICIOCCIO, R. A., TYRER, J., LOMAS, E.,

SHADFORTH, D., HOGDALL, E., HOGDALL, C., MCGUIRE, V., WHITTEMORE, A. S., EASTON, D. F., PONDER, B. A. J., KJAER, S. K., PHAROAH, P. D. P. & GAYTHER, S. A. 2006. Common variants in mismatch repair genes and risk of invasive ovarian cancer. *Carcinogenesis,* 27**,** 2235-2242.

SONG, H., RAMUS, S. J., SHADFORTH, D., QUAYE, L., KJAER, S. K., DICIOCCIO, R. A., DUNNING, A. M., HOGDALL, E., HOGDALL, C., WHITTEMORE, A. S., MCGUIRE, V., LESUEUR, F., EASTON, D. F., JACOBS, I. J., PONDER, B. A. J., GAYTHER, S. A. & PHAROAH, P. D. P. 2006. Common Variants in RB1 Gene and Risk of Invasive Ovarian Cancer. *Cancer Research,* 66**,** 10220-10226.

SONG, H., RAMUS, S. J., TYRER, J., BOLTON, K. L., GENTRY-MAHARAJ, A., WOZNIAK, E., ANTON-CULVER, H., CHANG-CLAUDE, J., CRAMER, D. W., DICIOCCIO, R., DORK, T., GOODE, E. L., GOODMAN, M. T., SCHILDKRAUT, J. M., SELLERS, T., BAGLIETTO, L., BECKMANN, M. W., BEESLEY, J., BLAAKAER, J., CARNEY, M. E., CHANOCK, S., CHEN, Z., CUNNINGHAM, J. M., DICKS, E., DOHERTY, J. A., DURST, M., EKICI, A. B., FENSTERMACHER, D., FRIDLEY, B. L., GILES, G., GORE, M. E., DE VIVO, I., HILLEMANNS, P., HOGDALL, C., HOGDALL, E., IVERSEN, E. S., JACOBS, I. J., JAKUBOWSKA, A., LI, D., LISSOWSKA, J., LUBINSKI, J., LURIE, G., MCGUIRE, V., MCLAUGHLIN, J., MEDREK, K., MOORMAN, P. G., MOYSICH, K., NAROD, S., PHELAN, C., PYE, C., RISCH, H., RUNNEBAUM, I. B., SEVERI, G., SOUTHEY, M., STRAM, D. O., THIEL, F. C., TERRY, K. L., TSAI, Y.-Y., TWOROGER, S. S., VAN DEN BERG, D. J., VIERKANT, R. A., WANG-GOHRKE, S., WEBB, P. M., WILKENS, L. R., WU, A. H., YANG, H., BREWSTER, W., ZIOGAS, A., HOULSTON, R., TOMLINSON, I., WHITTEMORE, A. S., ROSSING, M. A., PONDER, B. A. J., PEARCE, C. L., NESS, R. B., MENON, U., KJAER, S. K., GRONWALD, J., GARCIA-CLOSAS, M., FASCHING, P. A., EASTON, D. F., CHENEVIX-TRENCH, G., BERCHUCK, A., PHAROAH, P. D. P. & GAYTHER, S. A. 2009. A genome-wide association study identifies a new ovarian cancer susceptibility locus on 9p22.2. *Nat Genet,* 41**,** 996-1000.

SONG, H. & XU, Y. 2007. Gain of function of p53 cancer mutants in disrupting critical DNA damage response pathways. *Cell Cycle,* 6**,** 1570 - 1573.

SONODA, E., TAKATA, M., YAMASHITA, Y. M., MORRISON, C. & TAKEDA, S. 2001. Homologous DNA recombination in vertebrate cells. *Proc Natl Acad Sci U S A,* 98**,** 8388-94.

SOOD, A., SOROSKY, J., DOLAN, M., ANDERSON, B. & BULLER, R. 1999. Distant metastases in ovarian cancer: association with p53 mutations. *Clin Cancer Res,* 5**,** 2485 - 2490.

SOSLOW, R. A. 2008. Histologic subtypes of ovarian carcinoma: an overview. *Int J Gynecol Pathol,* 27**,** 161-74.

SOUNG, Y. H., JEONG, E. G., AHN, C. H., KIM, S. S., SONG, S. Y., YOO, N. J. & LEE, S. H. 2008. Mutational analysis of caspase 1, 4, and 5 genes in common human cancers. *Hum Pathol,* 39**,** 895-900.

SPIELMAN, R., BASTONE, L., BURDICK, J., MORLEY, M., EWENS, W. & CHEUNG, V. 2007. Common genetic variants account for differences in gene expression among ethnic groups. *Nat Genet,* 39**,** 226 - 231.

SPURDLE, A. B., CHEN, X., ABBAZADEGAN, M., MARTIN, N., KHOO, S.-K., HURST, T., WARD, B., WEBB, P. M. & CHENEVIX-TRENCH, G. 2000. CYP17 promotor polymorphism and ovarian cancer risk. *International Journal of Cancer,* 86**,** 436-439.

STACEY, S., MANOLESCU, A., SULEM, P., RAFNAR, T., GUDMUNDSSON, J., GUDJONSSON, S., MASSON, G., JAKOBSDOTTIR, M., THORLACIUS, S., HELGASON, A., ABEN, K., STROBBE, L., ALBERS-AKKERS, M., SWINKELS, D., HENDERSON, B., KOLONEL, L., LE MARCHAND, L., MILLASTRE, E., ANDRES, R., GODINO, J., GARCIA-PRATS, M., POLO, E., TRES, A., MOUY, M., SAEMUNDSDOTTIR, J., BACKMAN, V., GUDMUNDSSON, L., KRISTJANSSON, K., BERGTHORSSON, J. & KOSTIC, J. 2007. Common variants on chromosomes 2q35 and 16q12 confer susceptibility to estrogen receptor-positive breast cancer. *Nat Genet,* 39**,** 865 - 869.

STACEY, S., MANOLESCU, A., SULEM, P., THORLACIUS, S., GUDJONSSON, S., JONSSON, G., JAKOBSDOTTIR, M., BERGTHORSSON, J., GUDMUNDSSON, J., ABEN, K., STROBBE, L., SWINKELS, D., VAN ENGELENBURG, K., HENDERSON, B., KOLONEL, L., LE MARCHAND, L., MILLASTRE, E., ANDRES, R., SAEZ, B., LAMBEA, J., GODINO, J., POLO, E., TRES, A., PICELLI, S., RANTALA, J.,

MARGOLIN, S., JONSSON, T., SIGURDSSON, H., JONSDOTTIR, T. & HRAFNKELSSON, J. 2008. Common variants on chromosome 5p12 confer susceptibility to estrogen receptor-positive breast cancer. *Nat Genet,* 40**,** 703 - 706.

STASIAK, A. Z., LARQUET, E., STASIAK, A., MULLER, S., ENGEL, A., VAN DYCK, E., WEST, S. C. & EGELMAN, E. H. 2000. The human Rad52 protein exists as a heptameric ring. *Curr Biol,* 10**,** 337-40.

STEFFENSEN, K. D., WALDSTROM, M., GROVE, A., LUND, B., PALLISGARD, N. & JAKOBSEN, A. 2011. Improved classification of epithelial ovarian cancer: results of 3 danish cohorts. *Int J Gynecol Cancer,* 21**,** 1592-600.

STONE, B., SCHUMMER, M., PALEY, P. J., THOMPSON, L., STEWART, J., FORD, M., CRAWFORD, M., URBAN, N., O'BRIANT, K. & NELSON, B. H. 2003. Serologic analysis of ovarian tumor antigens reveals a bias toward antigens encoded on 17q. *International Journal of Cancer,* 104**,** 73-84.

STOREY, D. J., RUSH, R., STEWART, M., RYE, T., AL-NAFUSSI, A., WILLIAMS, A. R., SMYTH, J. F. & GABRA, H. 2008. Endometrioid epithelial ovarian cancer. *Cancer,* 112**,** 2211-2220.

STRANGER, B., FORREST, M., CLARK, A., MINICHIELLO, M., DEUTSCH, S., LYLE, R., HUNT, S., KAHL, B., ANTONARAKIS, S., TAVARE, S., DELOUKAS, P. & DERMITZAKIS, E. 2005. Genome-wide associations of gene expression variation in humans. *PLoS Genet,* 1**,** e78.

STRANGER, B., NICA, A., FORREST, M., DIMAS, A., BIRD, C., BEAZLEY, C., INGLE, C., DUNNING, M., FLICEK, P., KOLLER, D., MONTGOMERY, S., TAVARE, S., DELOUKAS, P. & DERMITZAKIS, E. 2007. Population genomics of human gene expression. *Nat Genet,* 39**,** 1217 - 1224.

STRATTON, J. F., GAYTHER, S. A., RUSSELL, P., DEARDEN, J., GORE, M., BLAKE, P., EASTON, D. & PONDER, B. A. 1997. Contribution of BRCA1 mutations to ovarian cancer. *N Engl J Med,* 336**,** 1125-30.

STRATTON, J. F., PHAROAH, P., SMITH, S. K., EASTON, D. & PONDER, B. A. 1998. A systematic review and meta-analysis of family history and risk of ovarian cancer. *Br J Obstet Gynaecol,* 105**,** 493-9.

STREULI, C. H., SCHMIDHAUSER, C., BAILEY, N., YURCHENCO, P., SKUBITZ, A. P., ROSKELLEY, C. & BISSELL, M. J. 1995. Laminin mediates tissue-specific gene expression in mammary epithelia. *J Cell Biol,* 129**,** 591-603.

SUNDFELDT, K. 2003. Cell-cell adhesion in the normal ovary and ovarian tumors of epithelial origin; an exception to the rule. *Molecular and Cellular Endocrinology,* 202**,** 89-96.

SUNDQVIST J, F. H., SEDDIGHZADEH M, VODOLAZKAIA A, FASSBENDER A, KYAMA C, BOKOR A, STEPHANSSON O, GEMZELL-DANIELSSON K, D'HOOGHE TM. 2011. Ovarian cancer-associated polymorphisms in the BNC2 gene among women with endometriosis. *Hum Reprod,* 26**,** 2253-7.

SUVASINI, R. & SOMASUNDARAM, K. 2010. Essential role of PI3-kinase pathway in p53-mediated transcription: Implications in cancer chemotherapy. *Oncogene,* 29**,** 3605-3618.

SWISHER, E. M., SAKAI, W., KARLAN, B. Y., WURZ, K., URBAN, N. & TANIGUCHI, T. 2008. Secondary BRCA1 Mutations in BRCA1-Mutated Ovarian Carcinomas with Platinum Resistance. *Cancer Research,* 68**,** 2581-2586.

TABUSE, M., OHTA, S., OHASHI, Y., FUKAYA, R., MISAWA, A., YOSHIDA, K., KAWASE, T., SAYA, H., THIRANT, C., CHNEIWEISS, H., MATSUZAKI, Y., OKANO, H., KAWAKAMI, Y. & TODA, M. 2011. Functional analysis of HOXD9 in human gliomas and glioma cancer stem cells. *Molecular Cancer,* 10**,** 60.

TAKATA, M., SASAKI, M. S., SONODA, E., MORRISON, C., HASHIMOTO, M., UTSUMI, H., YAMAGUCHI-IWAI, Y., SHINOHARA, A. & TAKEDA, S. 1998. Homologous recombination and non-homologous end-joining pathways of DNA double-strand break repair have overlapping roles in the maintenance of chromosomal integrity in vertebrate cells. *Embo J,* 17**,** 5497-508.

TAKETANI, T., TAKI, T., SHIBUYA, N., ITO, E., KITAZAWA, J., TERUI, K. & HAYASHI, Y. 2002. The HOXD11 Gene Is Fused to the NUP98 Gene in Acute Myeloid Leukemia with t(2;11)(q31;p15). *Cancer Research,* 62**,** 33-37.

TAKEUCHI, S., TAKEUCHI, N., FERMIN, A. C., TAGUCHI, H. & KOEFFLER, H. P. 2003. Frameshift mutations in caspase-5 and other target genes in leukemia and lymphoma cell lines having microsatellite instability. *Leukemia Research,* 27**,** 359-361.

TAN, D. S., ROTHERMUNDT, C., THOMAS, K., BANCROFT, E., EELES, R., SHANLEY, S., ARDERN-JONES, A., NORMAN, A., KAYE, S. B. & GORE, M. E. 2008. "BRCAness" syndrome in ovarian cancer: a case-control study describing the clinical features and outcome of patients with epithelial ovarian cancer associated with BRCA1 and BRCA2 mutations. *J Clin Oncol,* 26**,** 5530-6.

TAN, T. & CHU, G. 2002. p53 Binds and activates the xeroderma pigmentosum DDB2 gene in humans but not mice. *Mol Cell Biol,* 22**,** 3247 - 3254.

TAN, Y., ZHANG, B., WU, T., SKOGERBO, G. & ZHU, X. 2009. Transcriptional inhibiton of Hoxd4 expression by miRNA-10a in human breast cancer cells. *BMC molecular biology,* 10**,** 12.

TANAKA, Y., TERAI, Y., TANABE, A., SASAKI, H., SEKIJIMA, T., FUJIWARA, S., YAMASHITA, Y., KANEMURA, M., UEDA, M., SUGITA, M., FRANKLIN, W. A. & OHMICHI, M. 2007. Prognostic effect of epidermal growth factor receptor gene mutations and the aberrant phosphorylation of Akt and ERK in ovarian cancer. *Cancer Biol Ther,* 11**,** 50-7.

TAO, H., COX, D. & FRAZER, K. 2006. Allele-specific KRT1 expression is a complex trait. *PLoS Genet,* 2**,** e93.

TAYLOR, H. S., VANDEN HEUVEL, G. B. & IGARASHI, P. 1997. A conserved Hox axis in the mouse and human female reproductive system: late establishment and persistent adult expression of the Hoxa cluster genes. *Biology of Reproduction,* 57**,** 1338-1345.

TELENIUS, H. K., CARTER, N. P., BEBB, C. E., NORDENSKJO¨LD, M., PONDER, B. A. J. & TUNNACLIFFE, A. 1992. Degenerate oligonucleotide-primed PCR: General amplification of target DNA by a single degenerate primer. *Genomics,* 13**,** 718-725.

TENERIELLO, M., EBINA, M., LINNOILA, R., HENRY, M., NASH, J., PARK, R. & BIRRER, M. 1993. p53 and Ki-ras gene mutations in epithelial ovarian neoplasms. *Cancer Res,* 53**,** 3103 - 3108.

TEODORIDIS, J. M., HALL, J., MARSH, S., KANNALL, H. D., SMYTH, C., CURTO, J., SIDDIQUI, N., GABRA, H., MCLEOD, H. L., STRATHDEE, G. & BROWN, R. 2005. CpG island methylation of DNA damage response genes in advanced ovarian cancer. *Cancer Res,* 65**,** 8961-7.

TERASAWA, K., SAGAE, S., TOYOTA, M., TSUKADA, K., OGI, K., SATOH, A., MITA, H., IMAI, K., TOKINO, T. & KUDO, R. 2004. Epigenetic inactivation of TMS1/ASC in ovarian cancer. *Clin Cancer Res,* 10**,** 2000-6.

THE AUSTRALIAN OVARIAN CANCER STUDY GROUP/AUSTRALIAN CANCER, S. 2011. Integrated genomic analyses of ovarian carcinoma. *Nature,* 474**,** 609-615.

THE INTERNATIONAL HAPMAP, C. 2005. A haplotype map of the human genome. *Nature,* 437**,** 1299-1320.

THIERY, J. P. 2003. Epithelial-mesenchymal transitions in development and pathologies. *Curr Opin Cell Biol,* 15**,** 740-6.

THOMAS, G., JACOBS, K. B., YEAGER, M., KRAFT, P., WACHOLDER, S., ORR, N., YU, K., CHATTERJEE, N., WELCH, R., HUTCHINSON, A., CRENSHAW, A., CANCEL-TASSIN, G., STAATS, B. J., WANG, Z., GONZALEZ-BOSQUET, J., FANG, J., DENG, X., BERNDT, S. I., CALLE, E. E., FEIGELSON, H. S., THUN, M. J., RODRIGUEZ, C., ALBANES, D., VIRTAMO, J., WEINSTEIN, S., SCHUMACHER, F. R., GIOVANNUCCI, E., WILLETT, W. C., CUSSENOT, O., VALERI, A., ANDRIOLE, G. L., CRAWFORD, E. D., TUCKER, M., GERHARD, D. S., FRAUMENI, J. F., HOOVER, R., HAYES, R. B., HUNTER, D. J. & CHANOCK, S. J. 2008. Multiple loci identified in a genome-wide association study of prostate cancer. *Nat Genet,* 40**,** 310-315.

THOMAS, P. A., KIRSCHMANN, D. A., CERHAN, J. R., FOLBERG, R., SEFTOR, E. A., SELLERS, T. A. & HENDRIX, M. J. C. 1999. Association between Keratin and Vimentin Expression, Malignant Phenotype, and Survival in Postmenopausal Breast Cancer Patients. *Clinical Cancer Research,* 5**,** 2698-2703.

TOMLINSON, I. P., WEBB, E., CARVAJAL-CARMONA, L., BRODERICK, P., HOWARTH, K., PITTMAN, A. M., SPAIN, S., LUBBE, S., WALTHER, A., SULLIVAN, K., JAEGER, E., FIELDING, S., ROWAN, A., VIJAYAKRISHNAN, J., DOMINGO, E., CHANDLER, I., KEMP, Z., QURESHI, M., FARRINGTON, S. M., TENESA, A., PRENDERGAST, J. G., BARNETSON, R. A., PENEGAR, S., BARCLAY, E., WOOD, W., MARTIN, L., GORMAN, M., THOMAS, H., PETO, J., BISHOP, D. T., GRAY, R., MAHER, E. R., LUCASSEN, A., KERR, D., EVANS, D. G., SCHAFMAYER, C., BUCH, S., VOLZKE, H., HAMPE, J., SCHREIBER, S., JOHN, U., KOESSLER, T., PHAROAH, P., VAN WEZEL,

T., MORREAU, H., WIJNEN, J. T., HOPPER, J. L., SOUTHEY, M. C., GILES, G. G., SEVERI, G., CASTELLVI-BEL, S., RUIZ-PONTE, C., CARRACEDO, A., CASTELLS, A., FORSTI, A., HEMMINKI, K., VODICKA, P., NACCARATI, A., LIPTON, L., HO, J. W., CHENG, K. K., SHAM, P. C., LUK, J., AGUNDEZ, J. A., LADERO, J. M., DE LA HOYA, M., CALDES, T., NIITTYMAKI, I., TUUPANEN, S., KARHU, A., AALTONEN, L., CAZIER, J. B., CAMPBELL, H., DUNLOP, M. G. & HOULSTON, R. S. 2008. A genome-wide association study identifies colorectal cancer susceptibility loci on chromosomes 10p14 and 8q23.3. *Nat Genet,* 40**,** 623-30.

TONE, A. A., BEGLEY, H., SHARMA, M., MURPHY, J., ROSEN, B., BROWN, T. J. & SHAW, P. A. 2008. Gene expression profiles of luteal phase fallopian tube epithelium from BRCA mutation carriers resemble high-grade serous carcinoma. *Clin Cancer Res,* 14**,** 4067-78.

TONIN PN, H. T., RODIER F, BOSSOLASCO M, LEE PD, NOVAK J, MANDERSON EN, PROVENCHER D, MES-MASSON AM. 2001. Microarray analysis of gene expression mirrors the biology of an ovarian cancer model. *Oncogene.,* 20**,** 6617-26.

TROESTER, M., HERSCHKOWITZ, J., OH, D., HE, X., HOADLEY, K., BARBIER, C. & PEROU, C. 2006. Gene expression patterns associated with p53 status in breast cancer. *BMC Cancer,* 6**,** 276.

TUCCI R, C. M., MATIZONKAS-ANTONIO LF, DURAZZO M, PINTO JUNIOR DDOS S, NUNES FD. 2011. HOXB5 expression in oral squamous cell carcinoma. *J Appl Oral Sci. ,* 19**,** 125-9.

TULIN, Y. J. A. A. V. 2010. The roles of PARP1 in gene control and cell differentiation. *Curr Opin Genet Dev.,* 20**,** 512–518.

TUTEJA, N. & TUTEJA, R. 2001. Unraveling DNA repair in human: molecular mechanisms and consequences of repair defect. *Crit Rev Biochem Mol Biol,* 36**,** 261-90.

TWOROGER, S. S., FAIRFIELD, K. M., COLDITZ, G. A., ROSNER, B. A. & HANKINSON, S. E. 2007. Association of oral contraceptive use, other contraceptive methods, and infertility with ovarian cancer risk. *Am J Epidemiol,* 166**,** 894-901.

UDLER, M., MAIA, A., CEBRIAN, A., BROWN, C., GREENBERG, D., SHAH, M., CALDAS, C., DUNNING, A., EASTON, D., PONDER, B. & PHAROAH, P. 2007. Common germline genetic variation in antioxidant defense genes and survival after diagnosis of breast cancer. *J Clin Oncol,* 25**,** 3015 - 3023.

ULLRICH, O., REINHECKEL, T., SITTE, N., HASS, R., GRUNE, T. & DAVIES, K. J. 1999. Poly-ADP ribose polymerase activates nuclear proteasome to degrade oxidatively damaged histones. *Proc Natl Acad Sci U S A,* 96**,** 6223-8.

UMAR, A. & KUNKEL, T. A. 1996. DNA-replication fidelity, mismatch repair and genome instability in cancer cells. *Eur J Biochem,* 238**,** 297-307.

VALLE, L., SERENA-ACEDO, T., LIYANARACHCHI, S., HAMPEL, H., COMERAS, I., LI, Z., ZENG, Q., ZHANG, H., PENNISON, M., SADIM, M., PASCHE, B., TANNER, S. & DE LA CHAPELLE, A. 2008. Germline allele-specific expression of TGFBR1 confers an increased risk of colorectal cancer. *Science,* 321**,** 1361 - 1365.

VAN ZANDWIJK, N., MATHY, A., BOERRIGTER, L., RUIJTER, H., TIELEN, I., DE JONG, D., BAAS, P., BURGERS, S. & NEDERLOF, P. 2007. EGFR and KRAS mutations as criteria for treatment with tyrosine kinase inhibitors: retro- and prospective observations in non-small-cell lung cancer. *Ann Oncol,* 18**,** 99 - 103.

VANDERHYDEN, B., SHAW, T. & ETHIER, J.-F. 2003. Animal models of ovarian cancer. *Reproductive Biology and Endocrinology,* 1**,** 67.

VENKITARAMAN, A. R. 2009. Targeting the Molecular Defect in BRCA-Deficient Tumors for Cancer Therapy. *Cancer Cell,* 16**,** 89-90.

VERA, C., TAPIA, V., KOHAN, K., GABLER, F., FERREIRA, A., SELMAN, A., VEGA, M. & ROMERO, C. 2012. Nerve Growth Factor Induces the Expression of Chaperone Protein Calreticulin in Human Epithelial Ovarian Cells. *Horm Metab Res,* 44**,** 639-643.

VERLAAN, D., GE, B., GRUNDBERG, E., HOBERMAN, R., LAM, K., KOKA, V., DIAS, J., GURD, S., MARTIN, N., MALLMIN, H., NILSSON, O., HARMSEN, E., KWAN, T. & PASTINEN, T. 2008. Targeted screening of cis-regulatory variation in human haplotypes. *Genome Res,* 19**,** 118 - 127.

VEY, N., MOZZICONACCI, M.-J., GROULET-MARTINEC, A., DEBONO, S., FINETTI, P., CARBUCCIA, N., BEILLARD, E., DEVILARD, E., ARNOULET, C., COSO, D., SAINTY, D., XERRI, L., STOPPA, A.-M., LAFAGE-POCHITALOFF, M., NGUYEN, C., HOULGATTE, R., BLAISE, D., MARANINCHI, D., BIRG, F., BIRNBAUM, D. &

BERTUCCI, F. 2004. Identification of new classes among acute myelogenous leukaemias with normal karyotype using gene expression profiling. *Oncogene,* 23**,** 9381-9391.

VULTUR, A., VILLANUEVA, J. & HERLYN, M. 2011. Targeting BRAF in Advanced Melanoma: A First Step toward Manageable Disease. *Clinical Cancer Research,* 17**,** 1658-1663.

WALES, M. M., BIEL, M. A., EL DEIRY, W., NELKIN, B. D., ISSA, J. P., CAVENEE, W. K., KUERBITZ, S. J. & BAYLIN, S. B. 1995. p53 activates expression of HIC-1, a new candidate tumour suppressor gene on 17p13.3. *Nat Med,* 1**,** 570-7.

WANG, B., HUROV, K., HOFMANN, K. & ELLEDGE, S. J. 2009. NBA1, a new player in the Brca1 A complex, is required for DNA damage resistance and checkpoint control. *Genes & Development,* 23**,** 729-739.

WANG, D. & SADEE, W. 2006. Searching for polymorphisms that affect gene expression and mRNA processing: example ABCB1 (MDR1). *Aaps J,* 8**,** E515-20.

WANG, D. G., FAN, J.-B., SIAO, C.-J., BERNO, A., YOUNG, P., SAPOLSKY, R., GHANDOUR, G., PERKINS, N., WINCHESTER, E., SPENCER, J., KRUGLYAK, L., STEIN, L., HSIE, L., TOPALOGLOU, T., HUBBELL, E., ROBINSON, E., MITTMANN, M., MORRIS, M. S., SHEN, N., KILBURN, D., RIOUX, J., NUSBAUM, C., ROZEN, S., HUDSON, T. J., LIPSHUTZ, R., CHEE, M. & LANDER, E. S. 1998. Large-Scale Identification, Mapping, and Genotyping of Single-Nucleotide Polymorphisms in the Human Genome. *Science,* 280**,** 1077-1082.

WANG X, W. D. T. 2012. The ups and downs of DNA repair biomarkers for PARP inhibitor therapies. *Am J Cancer Res,* 1**,** 301-327.

WEI, C., WU, Q., VEGA, V., CHIU, K., NG, P., ZHANG, T., SHAHAB, A., YONG, H., FU, Y. & WENG, Z. 2006. A global map of p53 transcription-factor binding sites in the human genome. *Cell,* 124**,** 207 - 219.

WEI, L., XUE, T., WANG, J., CHEN, B., LEI, Y., HUANG, Y., WANG, H. & XIN, X. 2009. Roles of clusterin in progression, chemoresistance and metastasis of human ovarian cancer. *Int J Cancer,* 125**,** 791-806.

WEINKAUF, M., CHRISTOPEIT, M., HIDDEMANN, W. & DREYLING, M. 2007. Proteome- and microarray-based expression analysis of lymphoma cell lines identifies a p53-centered cluster of differentially expressed proteins in mantle cell and follicular lymphoma. *Electrophoresis,* 28**,** 4416 - 4426.

WELCSH, P. L. & KING, M.-C. 2001. BRCA1 and BRCA2 and the genetics of breast and ovarian cancer. *Human Molecular Genetics,* 10**,** 705-713.

WELSH, J. B., ZARRINKAR, P. P., SAPINOSO, L. M., KERN, S. G., BEHLING, C. A., MONK, B. J., LOCKHART, D. J., BURGER, R. A. & HAMPTON, G. M. 2001. Analysis of gene expression profiles in normal and neoplastic ovarian tissue samples identifies candidate molecular markers of epithelial ovarian cancer. *Proceedings of the National Academy of Sciences,* 98**,** 1176-1181.

WENTZENSEN N, B. A., JACOBS K, YANG HP, BERG CD, CAPORASO N, PETERS U, RAGARD L, BUYS SS, CHANOCK S, HARTGE P. 2011. Genetic variation on 9p22 is associated with abnormal ovarian ultrasound results in the Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial. *PLoS One,* 6.

WHEELER, J. M. D., BODMER, W. F. & MORTENSEN, N. J. M. 2000. DNA mismatch repair genes and colorectal cancer. *Gut,* 47**,** 148-153.

WIEGAND, K. C., SHAH, S. P., AL-AGHA, O. M., ZHAO, Y., TSE, K., ZENG, T., SENZ, J., MCCONECHY, M. K., ANGLESIO, M. S., KALLOGER, S. E., YANG, W., HERAVI-MOUSSAVI, A., GIULIANY, R., CHOW, C., FEE, J., ZAYED, A., PRENTICE, L., MELNYK, N., TURASHVILI, G., DELANEY, A. D., MADORE, J., YIP, S., MCPHERSON, A. W., HA, G., BELL, L., FEREDAY, S., TAM, A., GALLETTA, L., TONIN, P. N., PROVENCHER, D., MILLER, D., JONES, S. J. M., MOORE, R. A., MORIN, G. B., OLOUMI, A., BOYD, N., APARICIO, S. A., SHIH, I.-M., MES-MASSON, A.-M., BOWTELL, D. D., HIRST, M., GILKS, B., MARRA, M. A. & HUNTSMAN, D. G. 2010. ARID1A Mutations in Endometriosis-Associated Ovarian Carcinomas. *New England Journal of Medicine,* 363**,** 1532-1543.

WILCOX, C. B., BAYSAL, B. E., GALLION, H. H., STRANGE, M. A. & DELOIA, J. A. 2005. High-resolution methylation analysis of the BRCA1 promoter in ovarian tumors. *Cancer Genet Cytogenet,* 159**,** 114-22.

WILKINS, J., SOUTHAM, L., PRICE, A., MUSTAFA, Z., CARR, A. & LOUGHLIN, J. 2007. Extreme context specificity in differential allelic expression. *Hum Mol Genet,* 16**,** 537 -

546.
WILLIAMS, R., CHAN, E., COWLEY, M. & LITTLE, P. 2007. The influence of genetic variation on gene expression. *Genome Res,* 17**,** 1707 - 1716.
WOOD, M. A., MCMAHON, S. B. & COLE, M. D. 2000. An ATPase/helicase complex is an essential cofactor for oncogenic transformation by c-Myc. *Mol Cell,* 5**,** 321-30.
WOOSTER, R., NEUHAUSEN, S. L., MANGION, J., QUIRK, Y., FORD, D., COLLINS, N., NGUYEN, K., SEAL, S., TRAN, T., AVERILL, D. & ET AL. 1994. Localization of a breast cancer susceptibility gene, BRCA2, to chromosome 13q12-13. *Science,* 265**,** 2088-90.
WU, G., WEI, R., CHENG, E., NGO, B. & LEE, W.-H. 2009. Hec1 Contributes to Mitotic Centrosomal Microtubule Growth for Proper Spindle Assembly through Interaction with Hice1. *Molecular Biology of the Cell,* 20**,** 4686-4695.
WU, M., SOLER, D. R., ABBA, M. C., NUNEZ, M. I., BAER, R., HATZIS, C., LLOMBART-CUSSAC, A., LLOMBART-BOSCH, A. & ALDAZ, C. M. 2007. CtIP silencing as a novel mechanism of tamoxifen resistance in breast cancer. *Mol Cancer Res,* 5**,** 1285-95.
WU, M., XU, L.-G., SU, T., TIAN, Y., ZHAI, Z. & SHU, H.-B. 2004. AMID is a p53-inducible gene downregulated in tumors. *Oncogene,* 23**,** 6815-6819.
WU, Q., LOTHE, R., AHLQUIST, T., SILINS, I., TROPE, C., MICCI, F., NESLAND, J., SUO, Z. & LIND, G. 2007. DNA methylation profiling of ovarian carcinomas and their in vitro models identifies HOXA9, HOXB5, SCGB3A1, and CRABP1 as novel targets. *Molecular Cancer,* 6**,** 45.
WU, R., ZHAI, Y., FEARON, E. R. & CHO, K. R. 2001. Diverse Mechanisms of β-Catenin Deregulation in Ovarian Endometrioid Adenocarcinomas. *Cancer Research,* 61**,** 8247-8255.
WU, Y., MOSER, M., BAUTCH, V. L. & PATTERSON, C. 2003. HoxB5 Is an Upstream Transcriptional Switch for Differentiation of the Vascular Endothelium from Precursor Cells. *Molecular and Cellular Biology,* 23**,** 5680-5691.
XIE, C., GOU, M.-L., YI, T., DENG, H., LI, Z.-Y., LIU, P., QI, X.-R., HE, X., WEI, Y. & ZHAO, X. 2011. Efficient Inhibition of Ovarian Cancer by Truncation Mutant of FILIP1L Gene Delivered by Novel Biodegradable Cationic Heparin-Polyethyleneimine Nanogels. *Human Gene Therapy*.
XIONG, B., LI, S., AI, J.-S., YIN, S., OUYANG, Y.-C., SUN, S.-C., CHEN, D.-Y. & SUN, Q.-Y. 2008. BRCA1 Is Required for Meiotic Spindle Assembly and Spindle Assembly Checkpoint Activation in Mouse Oocytes. *Biology of Reproduction,* 79**,** 718-726.
YAMASHITA T, T. S., YAWEI Z, KATAYAMA H, KATO Y, NISHIWAKI K, YOKOHAMA Y, ISHIKAWA M. 2006. Suppression of invasive characteristics by antisense introduction of overexpressed HOX genes in ovarian cancer cells. *Int J Oncol.,* 28**,** 931-8.
YAN, H., YUAN, W., VELCULESCU, V., VOGELSTEIN, B. & KINZLER, K. 2002. Allelic variation in human gene expression. *Science,* 297**,** 1143.
YAN, H. & ZHOU, W. 2004. Allelic variations in gene expression. *Current opinion in oncology,* 16**,** 39 - 43.
YAN, W., LIU, G., SCOUMANNE, A. & CHEN, X. 2008. Suppression of inhibitor of differentiation 2, a target of mutant p53, is required for gain-of-function mutations. *Cancer Res,* 68**,** 6789 - 6796.
YANAI-INBAR, I. & SILVERBERG, S. G. 2000. Mucosal epithelial proliferation of the fallopian tube: prevalence, clinical associations, and optimal strategy for histopathologic assessment. *Int J Gynecol Pathol,* 19**,** 139-44.
YANCIK, R. 1993. Ovarian cancer. Age contrasts in incidence, histology, disease stage at diagnosis, and mortality. *Cancer,* 71**,** 517-23.
YANG, C. Q., CHAN, K. Y. K., NGAN, H. Y. S., KHOO, U. S., CHIU, P. M., CHAN, Q. K. Y., XUE, W. C. & CHEUNG, A. N. Y. 2006. Single nucleotide polymorphisms of follicle-stimulating hormone receptor are associated with ovarian cancer susceptibility. *Carcinogenesis,* 27**,** 1502-1506.
YANG, D., KHAN, S., SUN, Y., HESS, K., SHMULEVICH, I., SOOD, A. K. & ZHANG, W. 2011. Association of BRCA1 and BRCA2 mutations with survival, chemotherapy sensitivity, and gene mutator phenotype in patients with ovarian cancer. *Jama,* 306**,** 1557-65.
YANG, G., ROSEN, D. G., COLACINO, J. A., MERCADO-URIBE, I. & LIU, J. 2007a. Disruption of the retinoblastoma pathway by small interfering RNA and ectopic expression of the catalytic subunit of telomerase lead to immortalization of human ovarian surface epithelial cells. *Oncogene,* 26**,** 1492-8.

YANG, G., ROSEN, D. G., MERCADO-URIBE, I., COLACINO, J. A., MILLS, G. B., BAST, R. C., JR., ZHOU, C. & LIU, J. 2007b. Knockdown of p53 combined with expression of the catalytic subunit of telomerase is sufficient to immortalize primary human ovarian surface epithelial cells. *Carcinogenesis,* 28**,** 174-82.

YANG, Z., ZHANG, H. & KUMAR, R. 2005. Regulation of E-cadherin. *Breast Cancer Online,* 8**,** null-null.

YEMELYANOVA, A., VANG, R., KSHIRSAGAR, M., LU, D., MARKS, M. A., SHIH, I. M. & KURMAN, R. J. 2011. Immunohistochemical staining patterns of p53 can serve as a surrogate marker for TP53 mutations in ovarian carcinoma: an immunohistochemical and nucleotide sequencing analysis. *Mod Pathol,* 24**,** 1248-1253.

YODER, J. A., WALSH, C. P. & BESTOR, T. H. 1997. Cytosine methylation and the ecology of intragenomic parasites. *Trends Genet,* 13**,** 335-40.

YOON, J. H., DAMMANN, R. & PFEIFER, G. P. 2001. Hypermethylation of the CpG island of the RASSF1A gene in ovarian and renal cell carcinomas. *Int J Cancer,* 94**,** 212-7.

YOSHIDA, J., HORIUCHI, A., KIKUCHI, N., HAYASHI, A., OSADA, R., OHIRA, S., SHIOZAWA, T. & KONISHI, I. 2009. Changes in the expression of E-cadherin repressors, Snail, Slug, SIP1, and Twist, in the development and progression of ovarian carcinoma: the important role of Snail in ovarian tumorigenesis and progression. *Med Mol Morphol,* 42**,** 82-91.

YOU, L., CHANG, D., DU, H.-Z. & ZHAO, Y.-P. 2011. Genome-wide screen identifies PVT1 as a regulator of Gemcitabine sensitivity in human pancreatic cancer cells. *Biochemical and Biophysical Research Communications,* 407**,** 1-6.

YU, X. & BAER, R. 2000. Nuclear Localization and Cell Cycle-specific Expression of CtIP, a Protein That Associates with the BRCA1 Tumor Suppressor. *Journal of Biological Chemistry,* 275**,** 18541-18549.

YUASA, Y., NAGASAKI, H., OZE, I., AKIYAMA, Y., YOSHIDA, S., SHITARA, K., ITO, S., HOSONO, S., WATANABE, M., ITO, H., TANAKA, H., KANG, D., PAN, K. F., YOU, W. C. & MATSUO, K. 1002. Insulin-like growth factor 2 hypomethylation of blood leukocyte DNA is associated with gastric cancer risk. *Int J Cancer,* 2012**,** 27554.

ZAPATA, L., WHITEMAN, M., MARCHBANKS, P. & CURTIS, K. 2010. Intrauterine device use among women with ovarian cancer: a systematic review. *Contraception,* 82**,** 38-40.

ZAVADIL, J. & BOTTINGER, E. P. 2005. TGF-[beta] and epithelial-to-mesenchymal transitions. *Oncogene,* 24**,** 5764-5774.

ZHANG, L., CUI, X., SCHMITT, K., HUBERT, R., NAVIDI, W. & ARNHEIM, N. 1992. Whole genome amplification from a single cell: implications for genetic analysis. *Proc Natl Acad Sci U S A,* 89**,** 5847-51.

ZHENG, L., DAI, H., ZHOU, M., LI, X., LIU, C., GUO, Z., WU, X., WU, J., WANG, C., ZHONG, J., HUANG, Q., GARCIA-AGUILAR, J., PFEIFER, G. P. & SHEN, B. 2012. Polyploid cells rewire DNA damage response networks to overcome replication stress-induced barriers for tumour progression. *Nat Commun,* 3**,** 815.

ZHENG, W., LUO, F., LU, J. J., BALTAYAN, A., PRESS, M. F., ZHANG, Z. F. & PIKE, M. C. 2000. Reduction of BRCA1 expression in sporadic ovarian cancer. *Gynecol Oncol,* 76**,** 294-300.

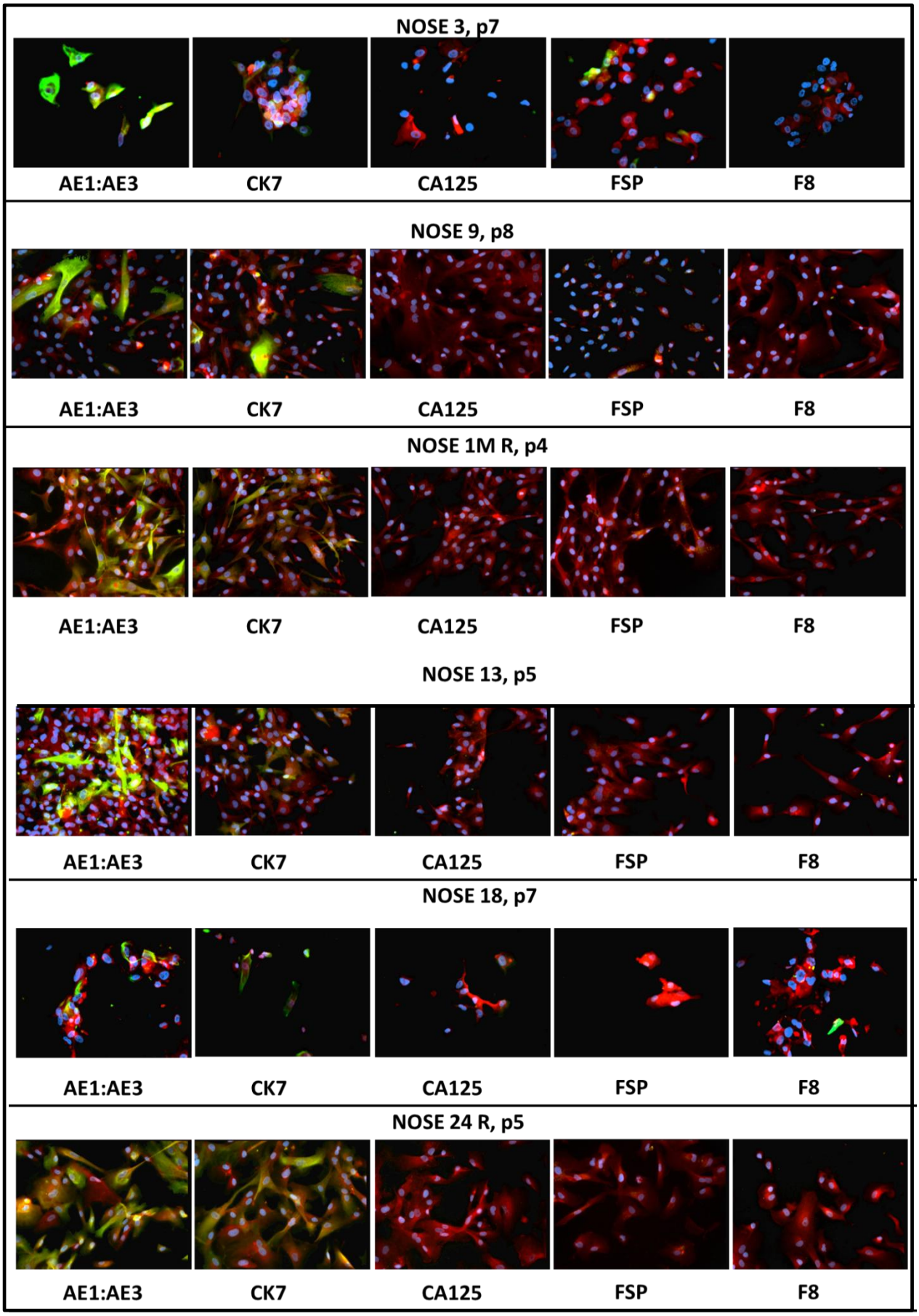
ZHU, C., DA CUNHA SANTOS, G., DING, K., SAKURADA, A., CUTZ, J., LIU, N., ZHANG, T., MARRANO, P., WHITEHEAD, M. & SQUIRE, J. 2008. Role of KRAS and EGFR as biomarkers of response to erlotinib in National Cancer Institute of Canada Clinical Trials Group Study BR.21. *J Clin Oncol,* 26**,** 4268 - 4275.

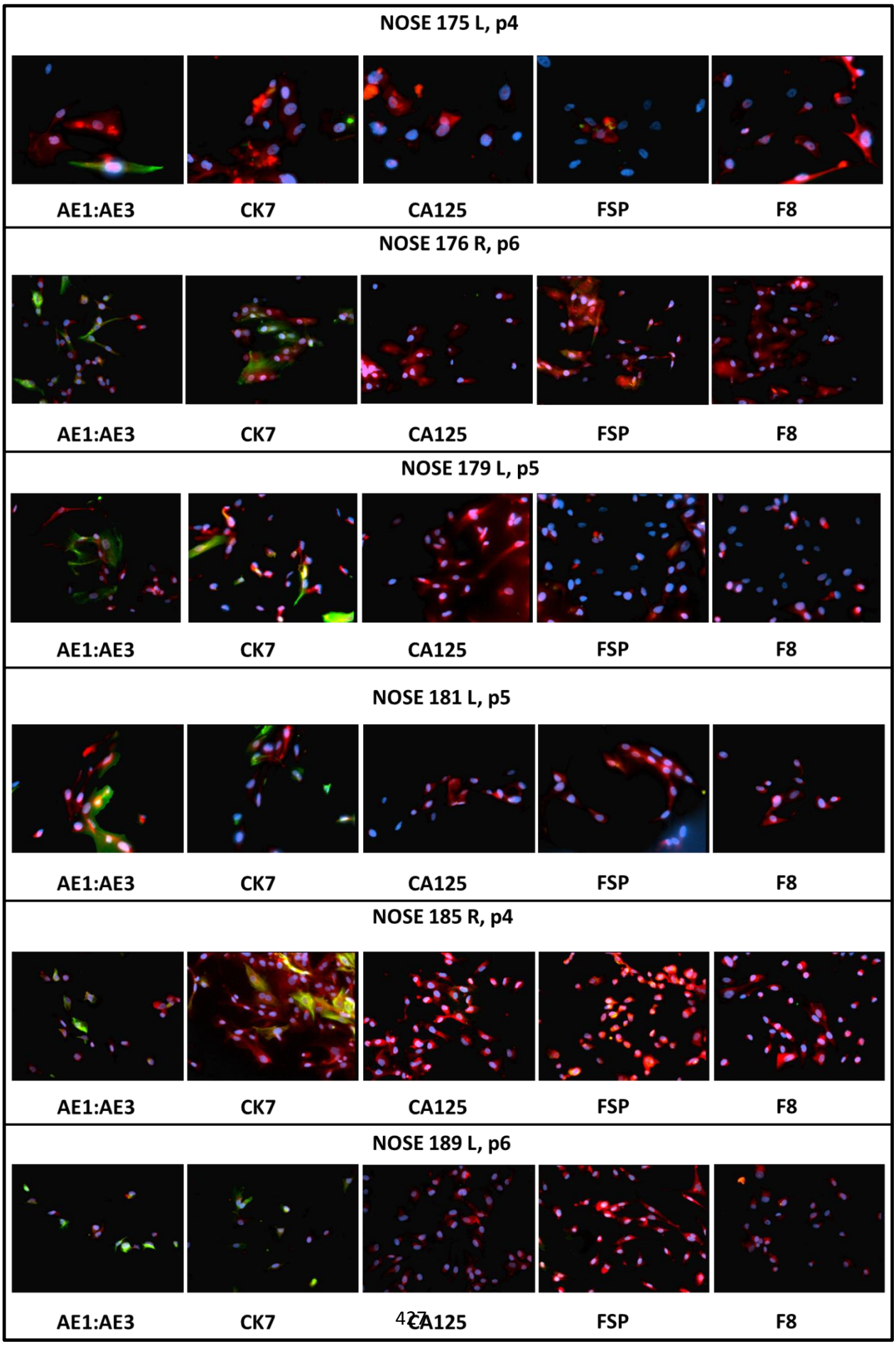
ZIETARSKA, M., MAUGARD, C. M., FILALI-MOUHIM, A., ALAM-FAHMY, M., TONIN, P. N., PROVENCHER, D. M. & MES-MASSON, A. M. 2007. Molecular description of a 3D in vitro model for the study of epithelial ovarian cancer (EOC). *Mol Carcinog,* 46**,** 872-85.

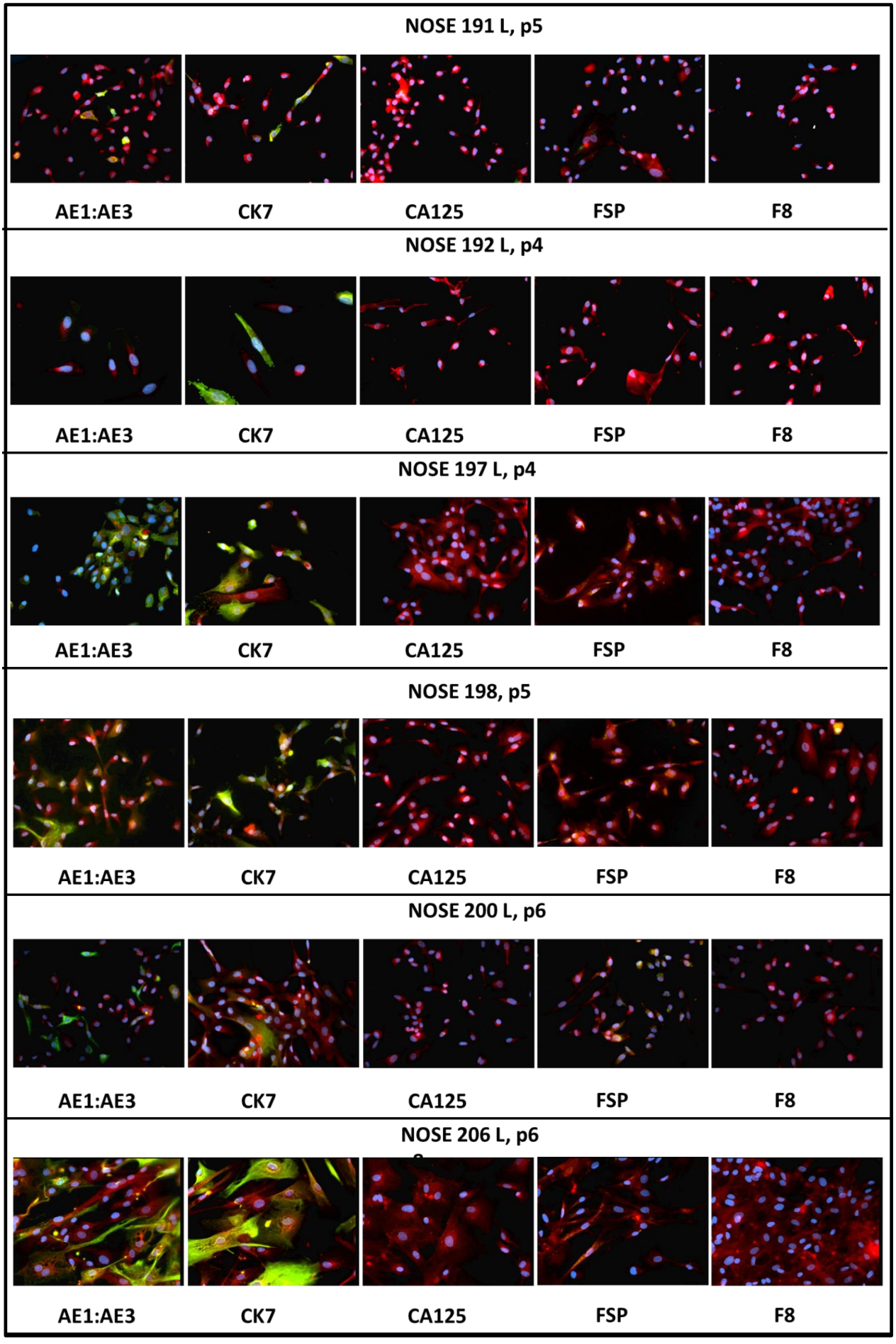
ZORN, K. K., JAZAERI, A. A., AWTREY, C. S., GARDNER, G. J., MOK, S. C., BOYD, J. & BIRRER, M. J. 2003. Choice of Normal Ovarian Control Influences Determination of Differentially Expressed Genes in Ovarian Cancer Expression Profiling Studies. *Clinical Cancer Research,* 9**,** 4811-4818.

Appendix 1: *NOSE & FTE in vitro models*

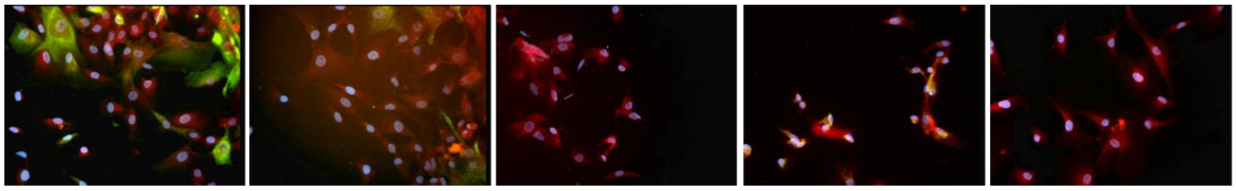
Figure 1: NOSE cell line expression of AE1:AE3, CK7, CA125, FSP, FVII as captured under a fluorescent microscope (Images at 200×).




NOSE 175 L, p4
AE1:AE3
CK7
CA125
FSP
F8
NOSE 176 R, p6
AE1:AE3
CK7
CA125
FSP
F8
NOSE 179 L, p5
AE1:AE3
CK7
CA125
FSP
F8
NOSE 181 L, p5
AE1:AE3
CK7
CA125
FSP
F8
NOSE 185 R, p4
AE1:AE3
CK7
CA125
FSP
F8
NOSE 189 L, p6
AE1:AE3
CK7
CA125
FSP
F8

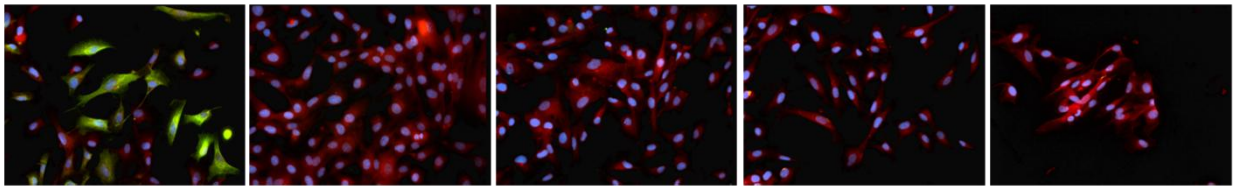

NOSE 191 L, p5
AE1:AE3
CK7
CA125
FSP
F8
NOSE 192 L, p4
AE1:AE3
CK7
CA125
FSP
F8
NOSE 197 L, p4
AE1:AE3
CK7
CA125
FSP
F8
NOSE 198, p5
AE1:AE3
CK7
CA125
FSP
F8
NOSE 200 L, p6
AE1:AE3
CK7
CA125
FSP
F8
NOSE 206 L, p6
AE1:AE3
CK7
CA125
FSP
F8

NOSE 211 L, p4



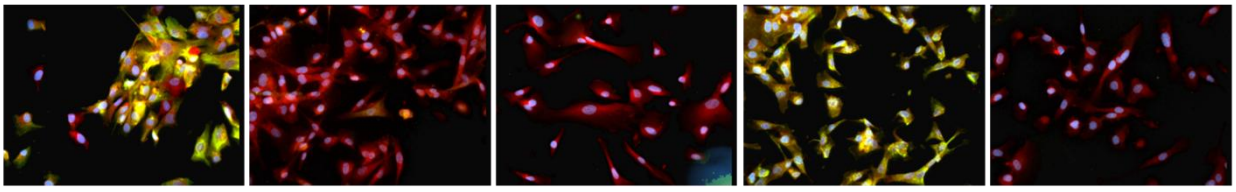
AE1:AE3 CK7 CA125 FSP F8

NOSE 216 L, p4



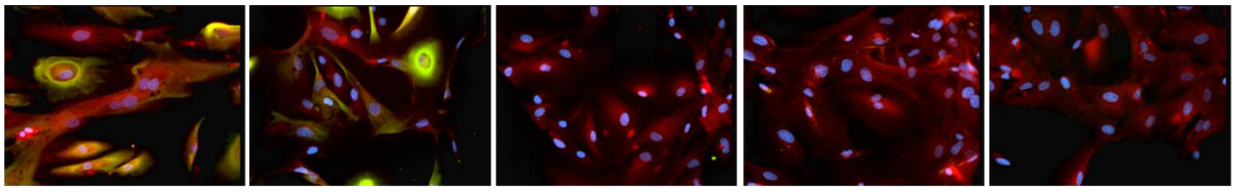
AE1:AE3 CK7 CA125 FSP F8

NOSE 217 L, p8



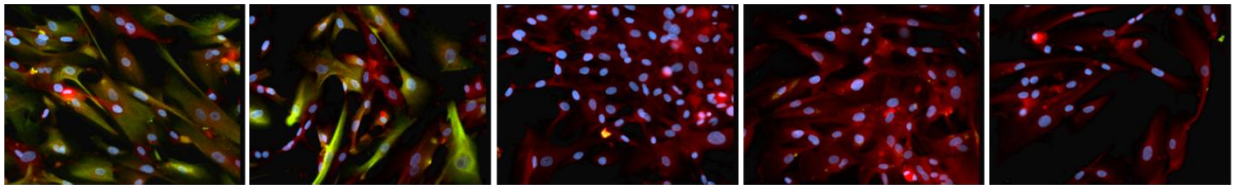
AE1:AE3 CK7 CA125 FSP F8

NOSE 218 L, p6



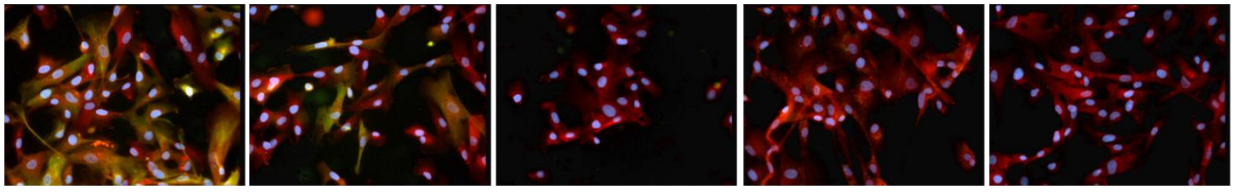
AE1:AE3 CK7 CA125 FSP F8

NOSE 224 L, p4

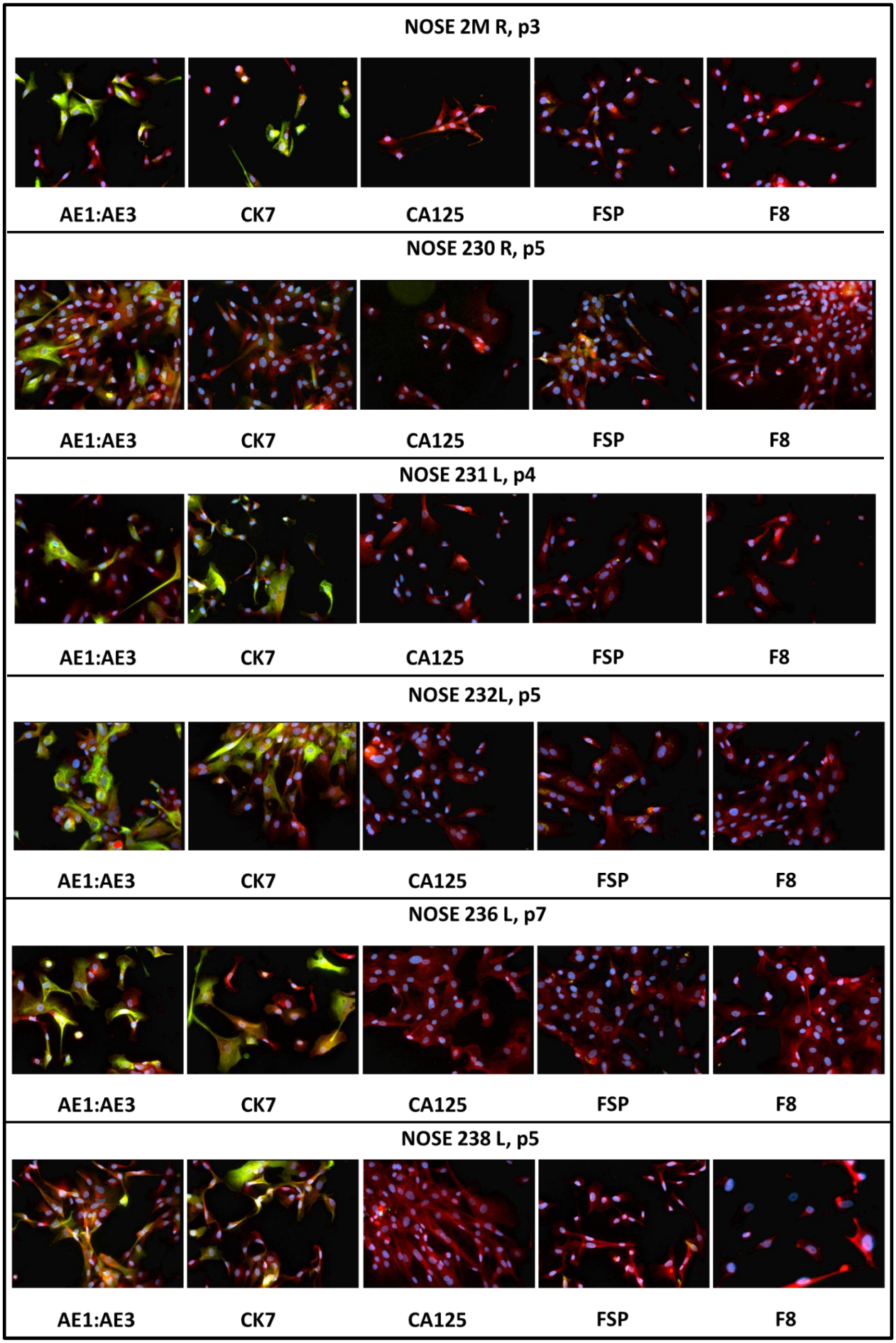


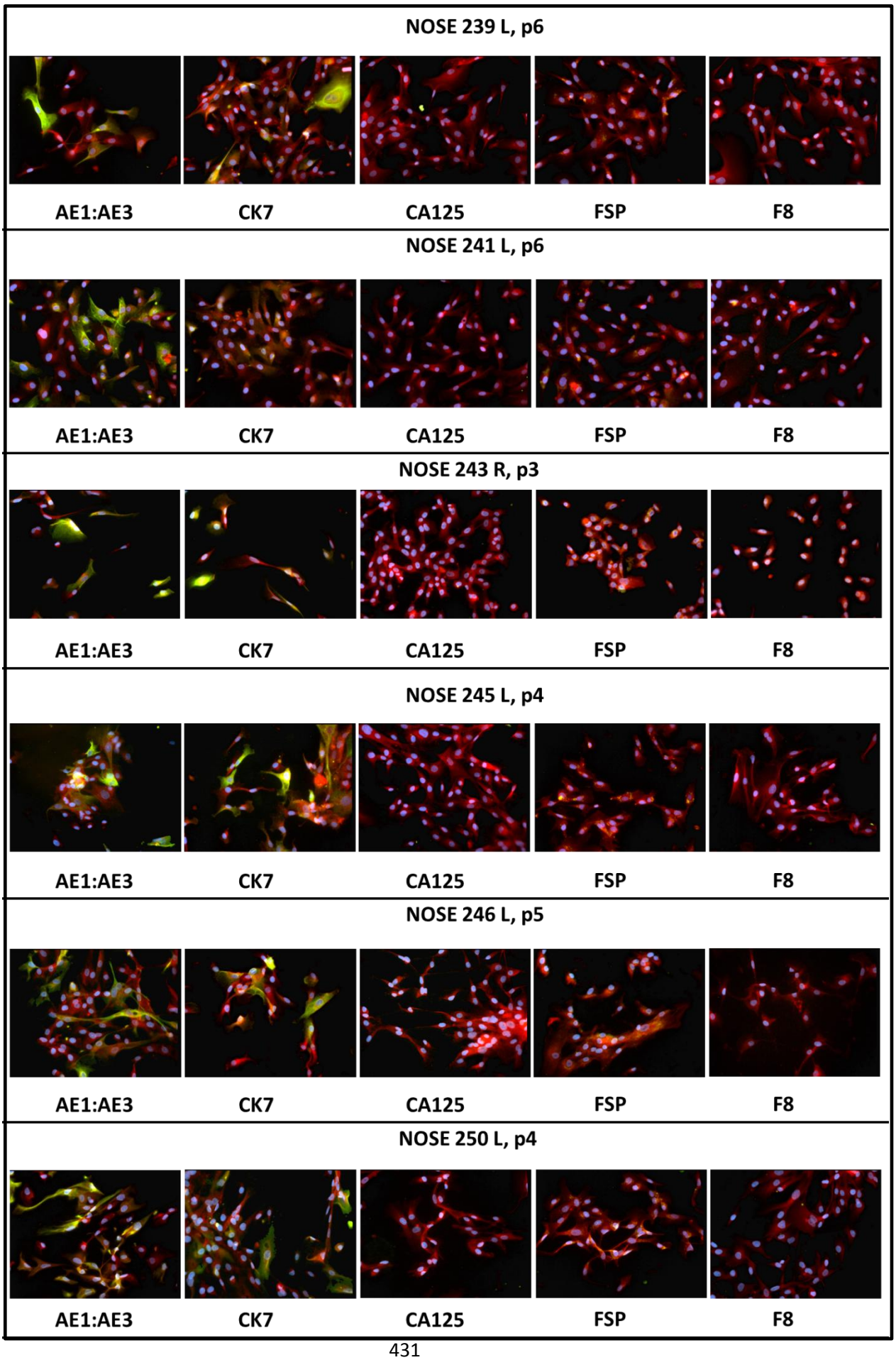
AE1:AE3 CK7 CA125 FSP F8

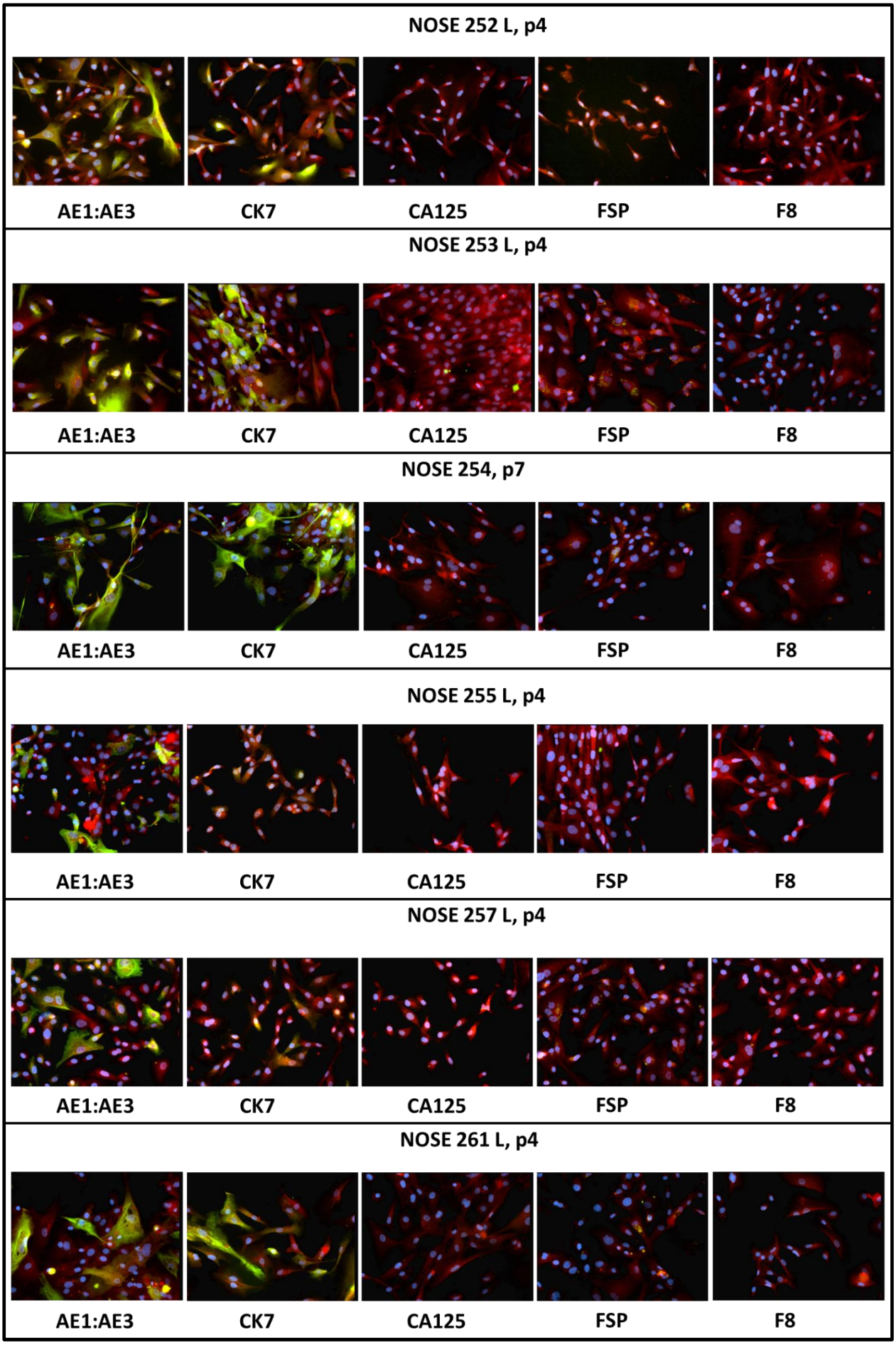
NOSE 228 R, p3

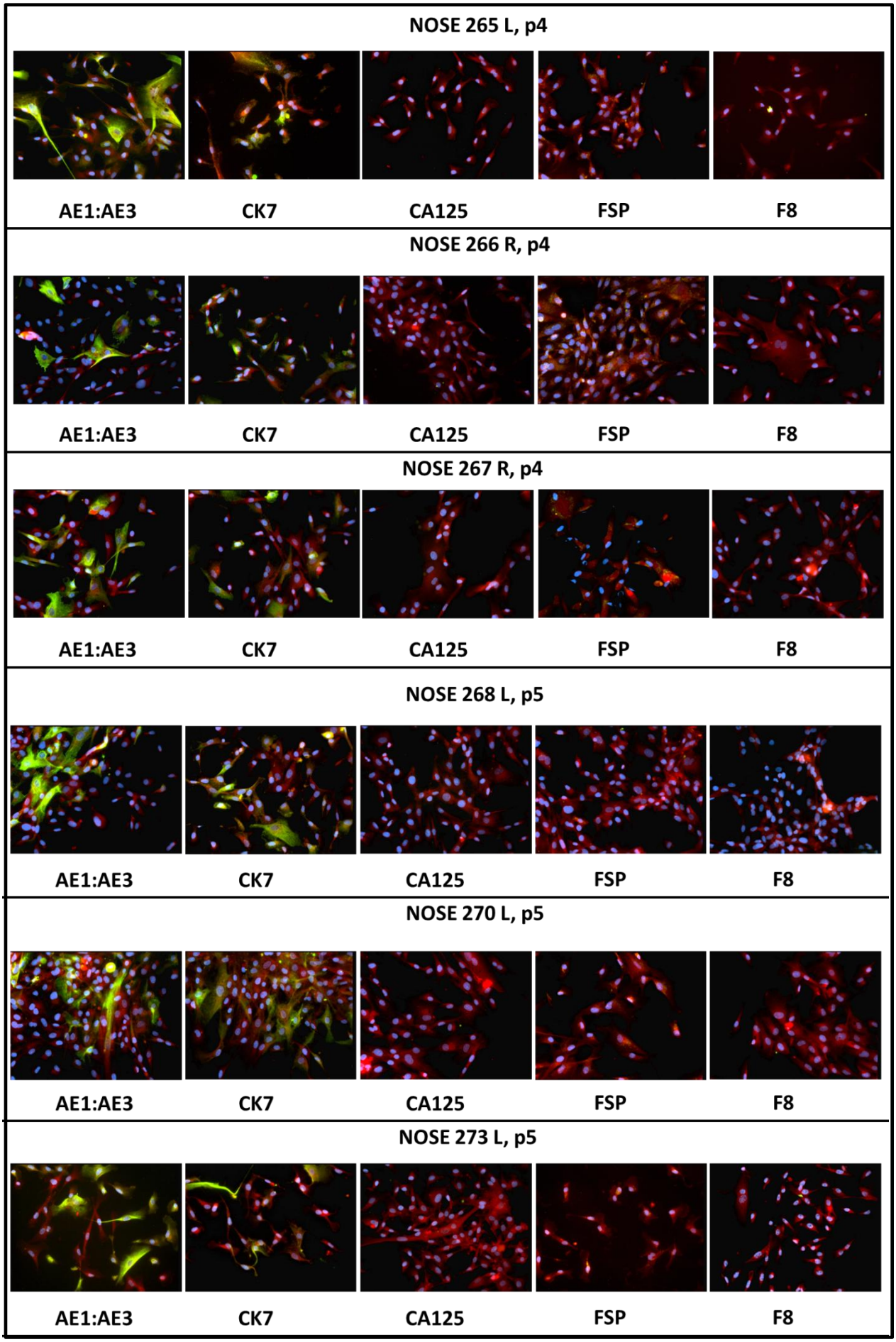


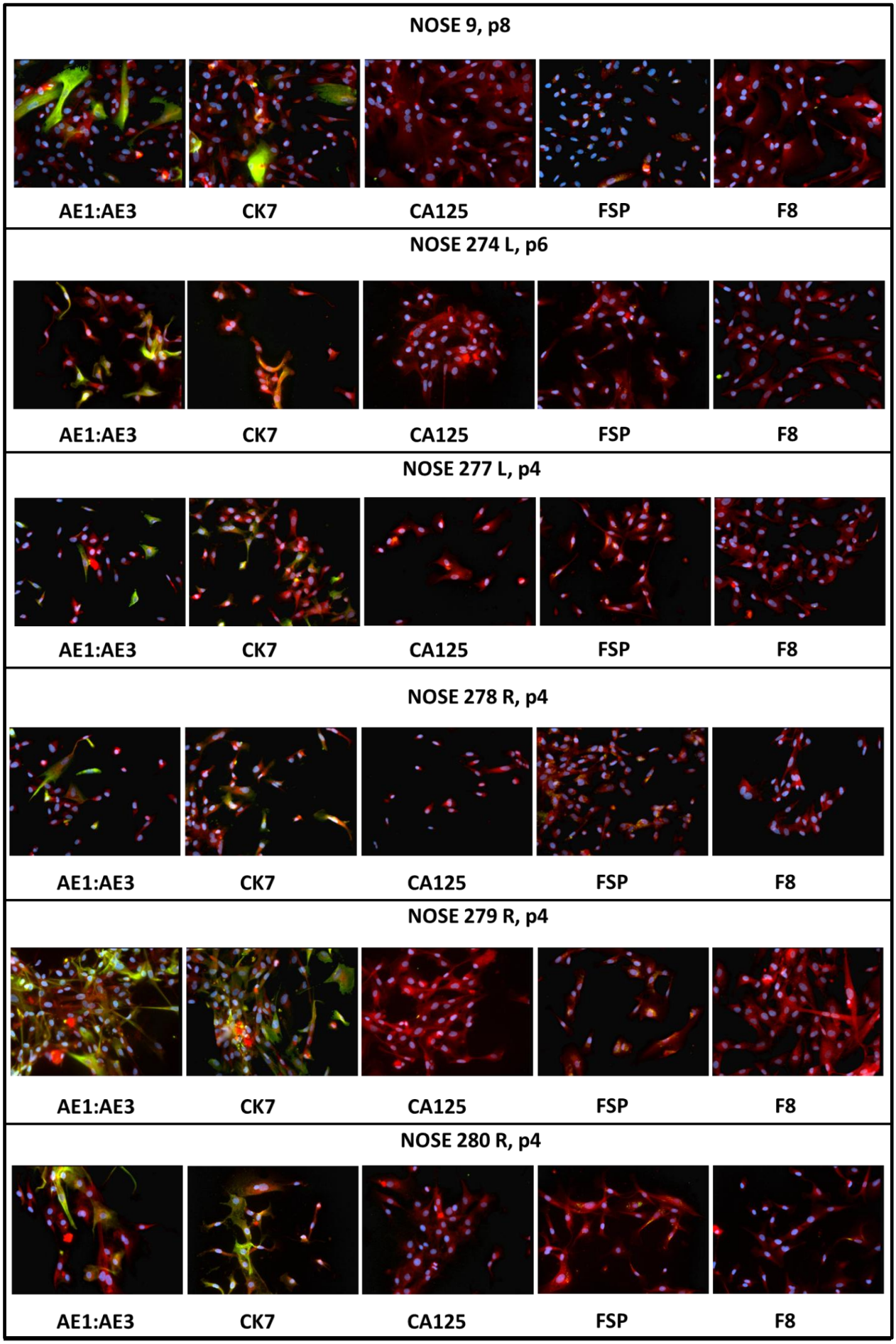
AE1:AE3 CK7 CA125 FSP F8


NOSE 2M R, p3
AE1:AE3
CK7
CA125
FSP
F8
NOSE 230 R, p5
AE1:AE3
CK7
CA125
FSP
F8
NOSE 231 L, p4
AE1:AE3
CK7
CA125
FSP
F8
NOSE 232L, p5
AE1:AE3
CK7
CA125
FSP
F8
NOSE 236 L, p7
AE1:AE3
CK7
CA125
FSP
F8
NOSE 238 L, p5
AE1:AE3
CK7
CA125
FSP
F8


NOSE 239 L, p6
AE1:AE3
CK7
CA125
FSP
F8
NOSE 241 L, p6
AE1:AE3
CK7
CA125
FSP
F8
NOSE 243 R, p3
AE1:AE3
CK7
CA125
FSP
F8
NOSE 245 L, p4
AE1:AE3
CK7
CA125
FSP
F8
NOSE 246 L, p5
AE1:AE3
CK7
CA125
FSP
F8
NOSE 250 L, p4
AE1:AE3
CK7
CA125
FSP
F8


NOSE 252 L, p4
AE1:AE3
CK7
CA125
FSP
F8
NOSE 253 L, p4
AE1:AE3
CK7
CA125
FSP
F8
NOSE 254, p7
AE1:AE3
CK7
CA125
FSP
F8
NOSE 255 L, p4
AE1:AE3
CK7
CA125
FSP
F8
NOSE 257 L, p4
AE1:AE3
CK7
CA125
FSP
F8
NOSE 261 L, p4
AE1:AE3
CK7
CA125
FSP
F8


NOSE 265 L, p4
AE1:AE3
CK7
CA125
FSP
F8
NOSE 266 R, p4
AE1:AE3
CK7
CA125
FSP
F8
NOSE 267 R, p4
AE1:AE3
CK7
CA125
FSP
F8
NOSE 268 L, p5
AE1:AE3
CK7
CA125
FSP
F8
NOSE 270 L, p5
AE1:AE3
CK7
CA125
FSP
F8
NOSE 273 L, p5
AE1:AE3
CK7
CA125
FSP
F8


NOSE 9, p8
AE1:AE3
CK7
CA125
FSP
F8
NOSE 274 L, p6
AE1:AE3
CK7
CA125
FSP
F8
NOSE 277 L, p4
AE1:AE3
CK7
CA125
FSP
F8
NOSE 278 R, p4
AE1:AE3
CK7
CA125
FSP
F8
NOSE 279 R, p4
AE1:AE3
CK7
CA125
FSP
F8
NOSE 280 R, p4
AE1:AE3
CK7
CA125
FSP
F8

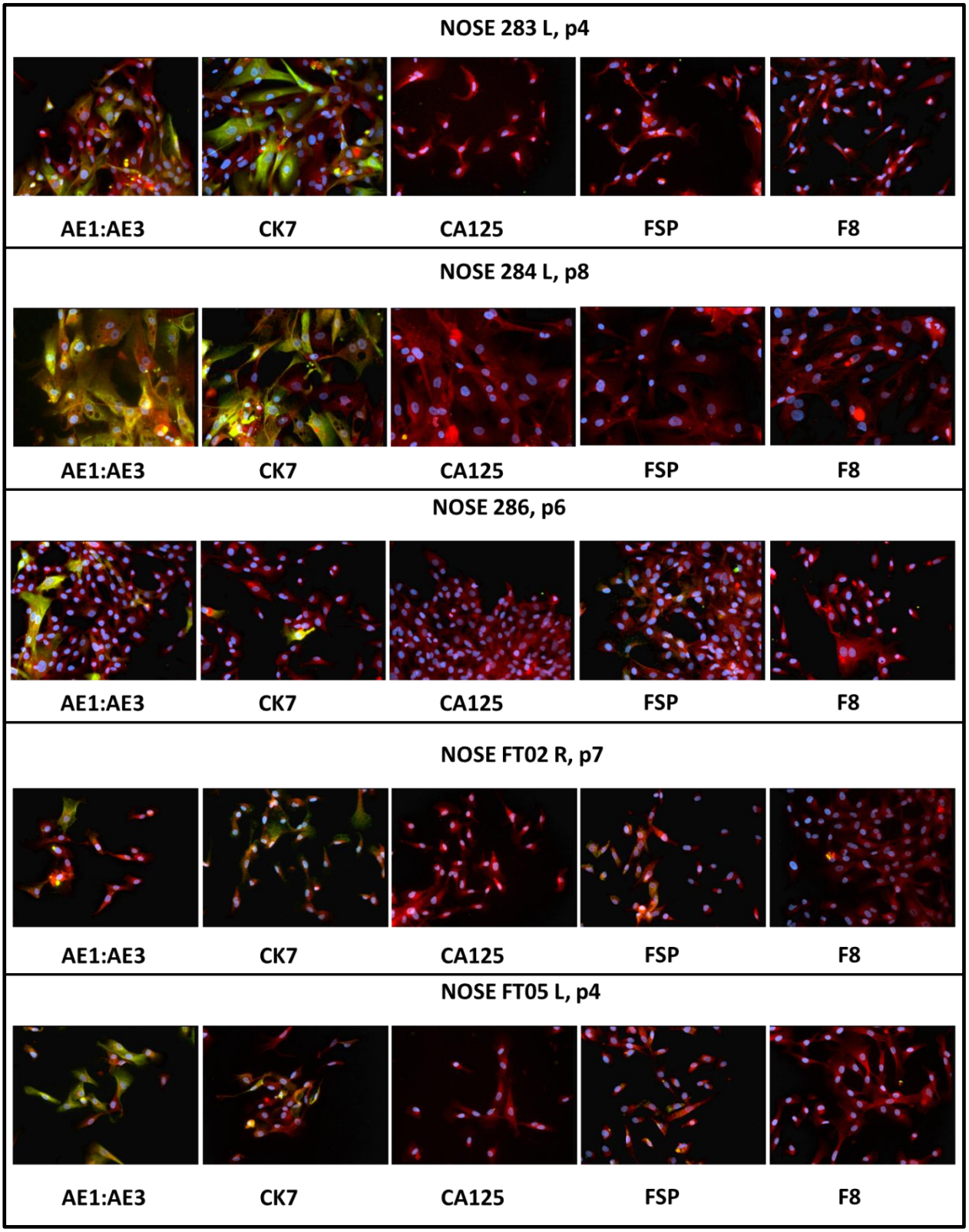

NOSE 283 L, p4
AE1:AE3
CK7
CA125
FSP
F8
NOSE 284 L, p8
AE1:AE3
CK7
CA125
FSP
F8
NOSE 286, p6
AE1:AE3
CK7
CA125
FSP
F8
NOSE FT02 R, p7
AE1:AE3
CK7
CA125
FSP
F8
NOSE FT05 L, p4
AE1:AE3
CK7
CA125
FSP
F8

Table 1: NOSE primary cell line collection with patient information such as age, diagnosis, histopathological examination of samples is presented. Ovarian surface epithelium brushings (OSE) were confirmed as normal (NOSE) after reported free of tumour by a pathologist. Samples highlighted in red were not taken forward due to malignancy. Samples highlighted in yellow were lost because of bacterial contamination. Sample highlighted in grey were not grown. (N/A= not applicable, where brushing was not collected).

Patient Number	Patient age	Histopatholgical diagnosis	Brushing Collected	Histopathology report (Right ovary)	Histopathology report (Left Ovary)
3	62	Grade 1 endometrioid endometrial carcinoma	(R) OSE	Free of tumour	N/A
9	62	Grade 2 endometrioid adenocarcinoma.	(R) OSE	Free of tumour	N/A
11	49	Severe complex atypical hyperplasia of endometrium.	OSE	Free of tumour	Free of tumour
13	66	Endometrioid clear cell carcinoma	OSE	Free of tumour	Free of tumour
18	62	Not specified	(L) OSE	Free of tumour	Free of tumour
24	69	Grade 2 endometrioid endometrial adenocarcinoma	(R) OSE	Free of tumour	N/A
175	74	High grade carcinosarcoma	(R), (L) OSE	Free of tumour	Free of tumour
176	59	Grade 1 endometrial adenocarcinoma, FIGO stage 1B	(L) OSE	Free of tumour	Free of tumour
179	68	Grade 2 endometrial serous papillary adenocarcinoma	(R), (L) OSE	Free of tumour	Free of tumour
181	61	Grade 2 endometrial endometriod adenocarcinoma	(R), (L) OSE	Free of tumour	Free of tumour
185	81	Grade 2 endometrioid endometrial adenocarcinoma	(R) OSE	Free of tumour	N/A
189	46	Grade 1 endometrioid adenocarcinoma	(R), (L) OSE	Free of tumour	Free of tumour
190	76	Grade 3 endometrial carcinoma	(R), (L) OSE	Free of tumour	Free of tumour
191	79	Moderate-poorly differentiated adenocarcinoma	(R), (L) OSE	Free of tumour	Free of tumour
192	82	High Grade serous endometrial adenocarcinoma	(R), (L) OSE	TUMOUR deposit	TUMOUR deposit
194	56	Grade 1, endometrioid adenocarcinoma	(R), (L) OSE	Free of tumour	Free of tumour
197	88	Grade 2 endometrioid endometrial adenocarcinoma	(R), (L) OSE	Free of tumour	Free of tumour
198	81	Grade 2 endometrial endometrioid adenocarcinoma	(R), (L) OSE	Free of tumour	Free of tumour
200	84	Grade 2 endometrioid adenocarcinoma	(R), (L) OSE	Free of tumour	Free of tumour
205	86	Grade 2 endometrial endometrioid adenocarcinoma	(L) OSE	N/A	Free of tumour
206	56	Simple hyperplasia with not atypia.	(L) OSE	N/A	Free of tumour
208	49	Papillary serous ovarian cystadenoma	(L) OSE	N/A	Serous cystadenoma
211	83	Grade 3 mixed clear cell and endometrioid carcinoma	(L) OSE	N/A	Free of tumour
216	71	Grade 3 endometrial endometrioid adenocarcinoma	(R), (L) OSE	Free of tumour	Free of tumour
217	65	Grade 2 endometrial endometriod adenocarcinoma	(R), (L) OSE	Free of tumour	Free of tumour
218	67	Grade 3 mixed endometriod/clear cell adenocarcinoma	(R), (L) OSE	Free of tumour	Free of tumour
221	61	Grade 2 endometrioid endometrial adenocarcinoma	(R), (L) OSE	Free of tumour	Free of tumour
222	40	Grade 3 endometrioid, endometrial adenocarcinoma	(L) OSE	N/A	TUMOUR deposit
223	84	Grade 2 endometrioid adenocarcinoma	(R), (L) OSE	Free of tumour	Free of tumour
224	75	Grade 2 endometrial Endometrioid adenocarcinoma	(R), (L) OSE	Free of tumour	Free of tumour
225	59	Grade 2 endometrioid endometrial adenocarcinoma	(R) OSE	Free of tumour	N/A
226	63	Grade 2 endometrioid endometrial adenocarcinoma	(R), (L) OSE	Free of tumour	Free of tumour
228	62	Grade 2 endometrioid endometrial adenocarcinoma	(R), (L) OSE	Free of tumour	Free of tumour
229	70	Grade 2 endometrioid endometrial adenocarcinoma	(R), (L) OSE	Free of tumour	Free of tumour
230	59	Grade 2 endometrioid endometrial adenocarcinoma	(R), (L) OSE	Free of tumour	Free of tumour
231	74	Carcinoma with heterologous elements, FIGO 1B	(R), (L) OSE	Free of tumour	Free of tumour
232	62	Carcinosarcma with heterologous elements, FIGO 1B	(R), (L) OSE	Free of tumour	Free of tumour
236	62	Grade 2 endometrial Endometrioid adenocarcinoma	(R), (L) OSE	Free of tumour	Free of tumour
237	80	High grade endometrial carcinosarcoma	(R), (L) OSE	Free of tumour	Free of tumour
238	73	Grade 2 endometrioid endometrial adenocarcinoma	(R), (L) OSE	Free of tumour	Free of tumour
239	82	High grade papillary serous endometrial carcinoma	(R), (L) OSE	Free of tumour	Free of tumour
241	75	Grade 3 endometriod adenocarcinoma	(L) OSE	N/A	Free of tumour
243	57	Invasive polypoid endometrial stromal tumour	(R) OSE	Free of tumour	N/A
245	70	Grade 3 endometrioid carcinoma.	(R), (L) OSE	Free of tumour	Free of tumour
246	75	Grade 3 poorly differentiated adenocarcinoma	(R), (L) OSE	Free of tumour	Free of tumour
250	66	Grade 3 endometrioid endometrial adenocarcinoma	(R), (L) OSE	Free of tumour	Free of tumour
252	40	Grade 2 mixed mucinus and endometrioid adenocarcinoma	(R), (L) OSE	Free of tumour	Free of tumour
253	60	Grade 2 endometrioid adenocarcinoma	(R), (L) OSE	Free of tumour	Free of tumour
254	70	Carcinosarcoma with heterologous differentiation	(R), (L) OSE	Free of tumour	Free of tumour
255	67	Grade 3 poorly differentiated adenocarcinoma	(R), (L) OSE	Free of tumour	Free of tumour
257	78	Grade 2 Endometrial endometrioid adenocarcinoma	(L) OSE	N/A	Free of tumour
258	70	Grade 3 endometrioid endometrial adenocarcinoma	(R), (L) OSE	Free of tumour	Free of tumour
261	81	High Grade papillary serous carcinoma of endometrium	(R), (L) OSE	Metastatic tumour	Free of tumour
265	63	CIN (Cervival intraepithelial neoplasia) 3 of cervix.	(R), (L) OSE	Free of tumour	Free of tumour
266	70	Grade 2 endometrioid endometrial carcinoma.	(R), (L) OSE	Free of tumour	Free of tumour
267	70	Grade 2 endometrioid adenocarcinoma into parametrium	(R) OSE	Free of tumour	N/A
268	77	Grade 3 endometrioid endometrial adenocarcinoma	(L) OSE	N/A	Free of tumour
270	65	Atypical complex hyperplasia	(R), (L) OSE	Free of tumour	Free of tumour
273	57	Grade 2 Endometrial endometrioid adenocarcinoma	(R), (L) OSE	Free of tumour	Free of tumour
274	52	Grade 1 well differentiated endometrioid adenocarcinoma	(R), (L) OSE	Free of tumour	Free of tumour
277	54	Grade 3 poorly differentiated endometrioid adenocarcinoma	(R), (L) OSE	Free of tumour	Free of tumour
278	63	Grade 2 endometrioid adenocarcinoma	(R), (L) OSE	Free of tumour	Free of tumour
279	72	Grade 2 endometrioid adenocarcinoma	(R), (L) OSE	Free of tumour	Free of tumour
280	69	Grade 3 endometrioid endometrial adenocarcinoma	(R), (L) OSE	Free of tumour	Free of tumour
283	65	Grade 2 endometrioid endometrial adenocarcinoma	(R), (L) OSE	Free of tumour	Free of tumour
284	60	Grade 2 endometrioid endometrial adenocarcinoma	(R), (L) OSE	Free of tumour	Free of tumour
286	71	Grade 2 endometrioid endometrial adenocarcinoma	(R), (L) OSE	Free of tumour	Free of tumour
NOSE FT02	50	Leiomyomas in uterus	(R) OSE	Free of tumour	N/A
NOSE FT05	54	Autolysed endometrium. Benign leiomyomas	(R), (L) OSE	Free of tumour	Free of tumour
1M R	44	Benign leiomyomas and ademonyosis	(R) OSE	Free of tumour	N/A
2M R	47	Uterus with multiple fibroids.	(R) OSE	Free of tumour	N/A

Number of cells seeded			FT01, p8, 4wks	FT02, p8, 4wks	FT03, p5, 4wks	FT05, p5, 4wks	FT283, p8, 4wks	FT284, p8, 4wks	112D 1000 cells
5000	Number of colonies	Replicate 1	496	312	668	1740	328	660	608
		Replicate 2	564	352	596	1848	392	548	556
		Replicate 3	604	404	760	1712	424	612	592
		Average	555	356	675	1767	381	607	585
	Colony forming efficiency	CFE 1	10	6	13	35	7	13	61
		CFE 2	11	7	12	37	8	11	56
		CFE 3	12	8	15	34	8	12	59
		CFE average	11	7	13	35	8	12	59
10000	Number of colonies	Replicate 1	936	752	1384	3688	584	1192	
		Replicate 2	1140	708	1248	3592	612	1024	
		Replicate 3	1028	732	1420	3872	544	1216	
		Average	1035	731	1351	3717	580	1144	
	Colony forming efficiency	CFE 1	9	8	14	37	6	12	
		CFE 2	11	7	12	36	6	10	
		CFE 3	10	7	14	39	5	12	
		CFE average	10	7	14	37	6	11	
20000	Number of colonies	Replicate 1	1988	1204	2696	6608	836	2372	
		Replicate 2	2248	1156	2316	6848	924	2268	
		Replicate 3	2132	1408	2472	6968	848	2016	
		Average	2123	1256	2495	6808	869	2219	
	Colony forming efficiency	CFE 1	10	6	13	33	4	12	
		CFE 2	11	6	12	34	5	11	
		CFE 3	11	7	12	35	4	10	
		CFE average	11	6	12	34	4	11	

Table 2: Colony counts in soft agar for FT cell lines after 4 weeks of seeding 5000, 10000 and 20000 cells in 3 replicates each. The EOC cell line 112D was used as a positive control for the experiment. The colony forming efficiency for each replicate was calculated and averaged.

Table 3: Commercially available EOC cell lines cultured to extract RNA from. The subtypes of these cell lines are shown. However as the collection of these cell lines spans over several years the terminology of histopathological examination has varied. The term “unknown histology” is referring to cell lines that their subtypes is probably serous but collected long ago before the subtypes were well differentiated from pathologists.

EOC cell line	Subtype
A2780	Serous
A2780CP	Serous
COV318	Serous
FUOV1	Serous
HEY	Serous
HEY A8	Serous
MPSCI	Serous
OAW 42	Serous
OAW-42M	Serous
OVCA 429	Serous
OVCA 433	Serous
OVCAR 5	Serous
OVCAR-3	Serous
OVM215	Serous
PEO-14	Serous
SKOV 3	Serous
SKOV 3IP	Serous
UWB1.289	Serous
UWB1.289+BRCA1	Serous
1847	*serous
1847 AD	*serous
CAOV3	*serous
COV 413	*serous
COV 624	*serous
DOV 13	*serous
HOC 7	*serous
IGROV1	*serous
INTOV-2	*serous
JAMA-2	*serous
LK1	*serous
LK2	*serous
OAW-41M	*serous
OC316	*serous
OVCA X	*serous
OVCAR 8	*serous
OVCAR-10	*serous
OVP1	*serous
PXN94	*serous
COV 644	Mucinous
EFO 27	Mucinous
COV-434	Granulosa cell
C13	Endometrioid
OV2008	Endometrioid
OV-TRL-90T	Endometrioid
TOV112D	Endometrioid
ES-2	Clear cell
TOV 21G	Clear cell

Table 4: Genotyping information for the established primary NOSE cell lines for 7 SNPs

NOSE cell line	rs2072590	rs344008	rs10088218	rs3814113	rs9303542	rs2363956	rs8170
NOSE 3	TG	AG	GG	TT	AG	TG	TC
NOSE 9	GG	GG	GG	TC	AA	TG	TC
NOSE 11	GG	GG	GG	TC	AA	TG	TC
NOSE 13	TG	GG	GG	TC	AG	TG	CC
NOSE 18	GG	AG	GG	CC	AG	TT	TC
NOSE 24	TG	AG	AG	TC	AA	TT	CC
NOSE 175	TG	GG	GG	TT	AG	TG	TC
NOSE 176	TG	AG	GG	TT	AA	TT	CC
NOSE 179	TG	GG	GG	CC	AG	GG	CC
NOSE 181	TG	GG	GG	TT	AA	TG	CC
NOSE 185	TG	GG	GG	TT	AA	TG	TC
NOSE 189	TG	GG	GG	TC	AA	TG	CC
NOSE 191	GG	GG	GG	TT	AG	TG	CC
NOSE 192	No data avaliable						
NOSE 197	GG	GG	AG	TT	AA	GG	CC
NOSE 198	TG	GG	AG	TC	AA	TT	TC
NOSE 200	GG	GG	GG	TT	AG	TG	CC
NOSE 206	GG	GG	AG	TT	AA	TT	CC
NOSE 211	TG	GG	GG	TC	AG	TG	TC
NOSE 216	TG	GG	GG	TC	AG	TT	TC
NOSE 217	GG	GG	GG	TT	AA	TT	CC
NOSE 218	GG	GG	AG	TC	AA	TG	CC
NOSE 224	GG	GG	GG	TT	AG	TT	CC
NOSE 228	GG	GG	GG	CC	AA	TG	CC
NOSE 229	GG	GG	GG	TT	AA	TT	TT
NOSE 230	TT	GG	GG	TT	AA	GG	CC
NOSE 231	GG	GG	AG	TC	AG	GG	CC
NOSE 232	TG	GG	GG	TT	AA	TT	CC
NOSE 236	GG	GG	GG	TC	AA	TG	TC
NOSE 238	GG	GG	AA	TC	AG	GG	CC
NOSE 239	TG	GG	GG	TT	AG	TT	TC
NOSE 241	TG	GG	GG	TC	GG	TT	CC
NOSE 243	TG	AG	GG	TC	AG	TG	CC
NOSE 245	GG	GG	GG	TT	AG	TG	CC
NOSE 246	GG	GG	GG	TC	AA	TG	TC
NOSE 250	GG	GG	GG	CC	AA	GG	CC
NOSE 252	TT	GG	GG	TT	GG	TG	CC
NOSE 253	GG	GG	AG	TT	AG	TT	CC
NOSE 254	GG	GG	GG	TT	AG	TT	CC
NOSE 255	TG	GG	AA	TC	GG	GG	CC
NOSE 257	GG	GG	GG	TC	AG	TT	TC
NOSE 261	TG	GG	AG	CC	AA	TG	CC
NOSE 265	TG	GG	AG	TC	AA	TG	CC
NOSE 266	TG	GG	AG	CC	AA	TT	CC
NOSE 267	GG	GG	GG	TC	AA	TG	CC
NOSE 268	TG	AG	GG	CC	AG	TG	CC
NOSE 270	GG	GG	GG	TC	AA	TT	CC
NOSE 273	GG	GG	AG	TT	AG	TG	TC
NOSE 274	TG	AG	GG	TT	AA	TG	CC
NOSE 277	GG	GG	GG	TT	AG	GG	CC
NOSE 278	TG	GG	GG	CC	AA	TG	CC
NOSE 279	TG	GG	GG	TC	AG	TG	CC
NOSE 280	TG	GG	GG	TT	GG	TT	TC
NOSE 283	GG	GG	GG	TC	AA	TG	TC
NOSE 284	No data avaliable						
NOSE 286	TG	GG	GG	TT	AA	TG	CC
NOSE 1MR	GG	GG	GG	0	AG	TT	TC
NOSE 2MR	TG	AG	GG	TC	AA	TG	CC
NOSEFT02	No data avaliable						
NOSEFT05	No data avaliable						

Appendix 2: *Functional analysis of MMCT-19 genes and tNSPs*

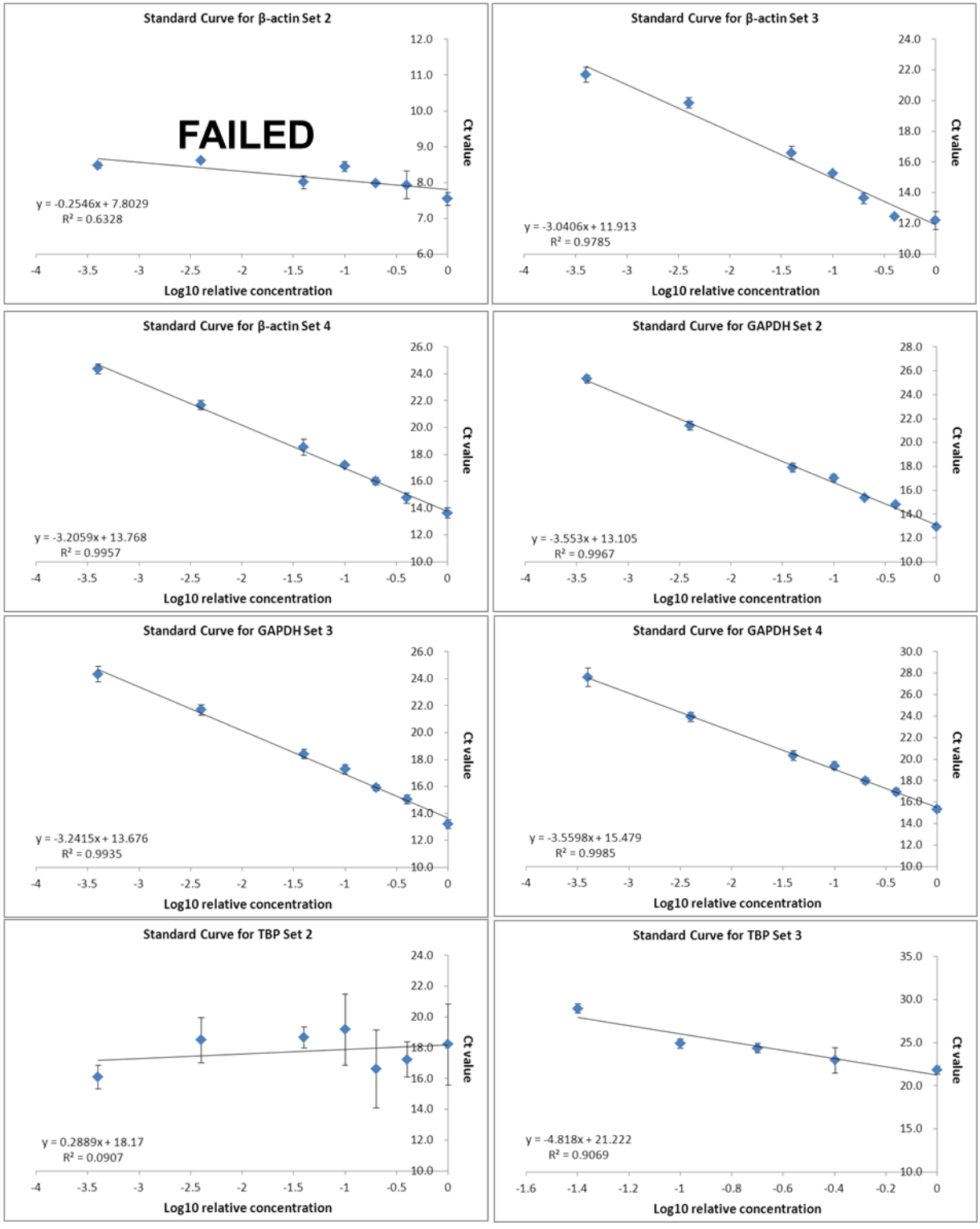
Table 1: Summary of quality control analysis for the SNPs selected to investigate allele specific LOH. Highlighted in red are the SNPs that failed. LOH= Loss of Heterozygosity. N/A: when analysis was not performed due to having failed sample QC or not having enough informative samples.

Gene	SNP	Order in gene	Plex	Number of samples	Genomic Call Rate %	% Genomic Concordances	Heterozygous duplicate samples	LOH in duplicate samples (Concordant/ discordant)	% of LOH concordance	Reason for excluding
AIFM2	rs2394655	1	28plex	301	100	100	1	0/1	0	LOH discordances
AIFM2	rs7908957	2	28plex	301	99	100	1	0/1	0	LOH discordances
AIFM2	rs1053495	3	33plex	239	100	100	9	8/1	89	PASSED
AIFM2	rs2271695	4	33plex	Non polymorphic						Non polymorphic
AIFM2	rs2894111	5	28plex	PCR failed						PCR failed
AIFM2	rs2394656	6	28plex	301	100	100	1	1/0	100	PASSED
AIFM2	rs6480440	7	33plex	239	96	100	26	25/1	96	PASSED
AIFM2	rs2280201	8	28plex	301	100	100	0	N/A	N/A	PASSED
AIFM2	rs10999147	9	33plex	239	100	100	15	15/0	100	PASSED
AIFM2	rs3750772	10	28plex	301	100	100	0	N/A	N/A	PASSED
AIFM2	rs4295944	11	33plex	239	100	100	39	26/13	67	LOH discordances
AIFM2	rs2394644	12	33plex	N/A	0	N/A				Genomic call rate
AIFM2	rs10999152	13	28plex	301	97	100	2	2/0	100	PASSED
AKTIP	rs9931702	1	28plex	301	100	100	3	3/0	100	PASSED
AKTIP	rs17801966	2	28plex	301	98	100	1	1/0	100	PASSED
AKTIP	rs7189819	3	28plex	301	99	100	7	3/4	0	PASSED
AKTIP	rs3743772	4	33plex	239	94	100	6	5/1	83	PASSED
AXIN2	rs11868547	1	28plex	301	99	100	2	2/0	100	PASSED
AXIN2	rs7591	2	33plex	239	100	100	44	41/3	95	PASSED
AXIN2	rs7210356	3	33plex	239	100	100	43	21/22	95	PASSED
AXIN2	rs11655966	4	28plex	N/A	0	N/A				Genomic call rate
AXIN2	rs4541111	5	28plex	301	100	85	N/A			Genomic discordances
AXIN2	rs4791171	6	33plex	239	100	100	29	28/1	97	PASSED
AXIN2	rs11079571	7	33plex	239	0	N/A				Genomic call rate
AXIN2	rs3923086	8	28plex	301	100	100	6	6/0	100	PASSED
CASP5	rs518604	1	33plex	239	99	100	63	60/3	95	PASSED
CASP5	rs523104	2	33plex	239	0	N/A				Genomic call rate
CASP5	rs3181328	3	28plex	301	100	100	0	N/A	N/A	PASSED
CASP5	rs17446518	4	28plex	301	99	77	N/A			Genomic discordances
CASP5	rs9651713	5	28plex	301	100	100	1	1/0	100	PASSED
CASP5	rs3181175	6	33plex	239	100	100	0	N/A	N/A	PASSED
CASP5	rs3181174	7	33plex	239	100	100	14	11/3	78	LOH discordances
CASP5	rs2282657	8	33plex	239	100	100	0	N/A	N/A	PASSED
FILIP1L	rs796977	1	33plex	239	99	100	40	40/0	98	PASSED
FILIP1L	rs793477	2	28plex	301	99	100	0	N/A	N/A	PASSED
FILIP1L	rs793446	3	33plex	239	100	100	41	29/12	71	LOH discordances
FILIP1L	rs17338680	4	33plex	239	100	69	N/A			Genomic discordances
FILIP1L	rs9864437	5	33plex	239	100	100	27	25/2	93	PASSED
FILIP1L	rs6788750	6	28plex	301	100	100	4	4/0	100	PASSED
FILIP1L	rs12494994	7	33plex	239	99	100	22	21/1	95	PASSED
RBBP8	rs7239066	1	28plex	301	100	100	0	N/A	N/A	PASSED
RBBP8	rs11082221	2	28plex	301	100	100	0	N/A	N/A	PASSED
RBBP8	rs9304261	3	33plex	239	99	100	0	N/A	N/A	PASSED
RGC32	rs3783194	1	28plex	239	76	N/A				Genomic call rate
RGC32	rs11618371	2	33plex	239	100	100	25	22/3	88	PASSED
RGC32	rs995845	3	33plex	239	100	100	31	30/1	97	PASSED
RGC32	rs975590	4	28plex	301	100	100	3	3/0	100	PASSED
RGC32	rs3783197	5	33plex	239	100	70	N/A			Genomic discordances
RUVBL1	rs9860614	1	33plex	239	99	100	4	3/1	75	LOH discordances
RUVBL1	rs13063604	2	28plex	301	76	69	N/A			Genomic call rate
RUVBL1	rs3732402	3	33plex	239	100	100	46	38/8	83	PASSED
RUVBL1	rs13091198	4	33plex	239	93	100	5	3/2	60	LOH discordances
RUVBL1	rs7650365	5	28plex	301	78	N/A				Genomic call rate
RUVBL1	rs4857836	6	33plex	239	100	100	37	35/2	95	PASSED
RUVBL1	rs9821568	7	28plex	301	100	100	2	2/0	100	PASSED
STAG3	rs11762932	1	33plex	239	100	100	29	27/2	96	PASSED
STAG3	rs2246713	2	28plex	301	100	77	N/A			Genomic discordances
STAG3	rs1637001	3	33plex	239	100	100	25	24/1	99	PASSED
BRCA1	rs799917		28plex	301	97	100	4	3/1	75	LOH discordances
BRCA2	rs144848		28plex	301	100	100	3	2/1	67	LOH discordances

Appendix 3: *Post-GWAS characterisation of risk loci*

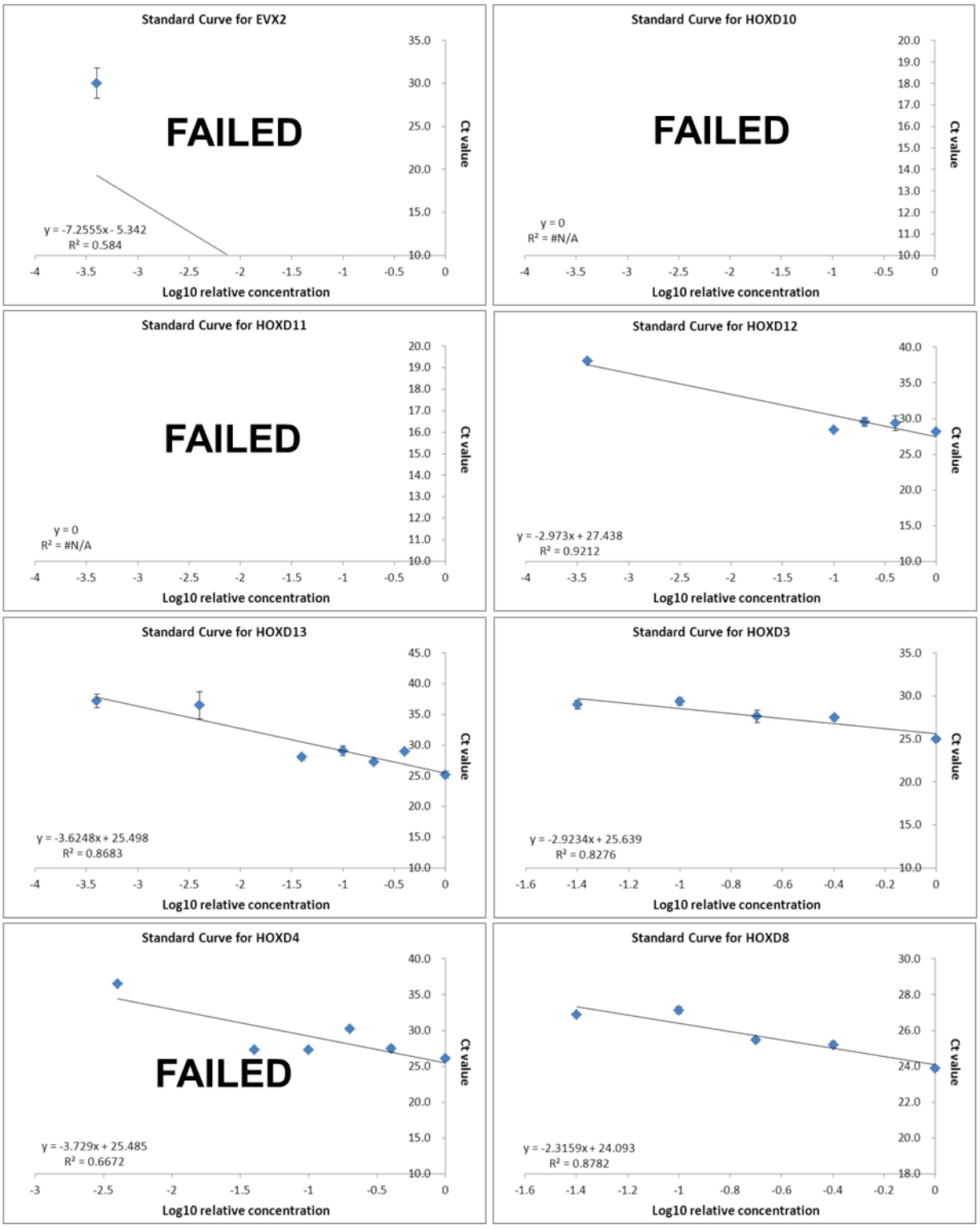
Figure 1: Standard Curves for assays run in the Fluidigm assay for gene expression analysis between EOC and normal (NOSE & FT) cell lines

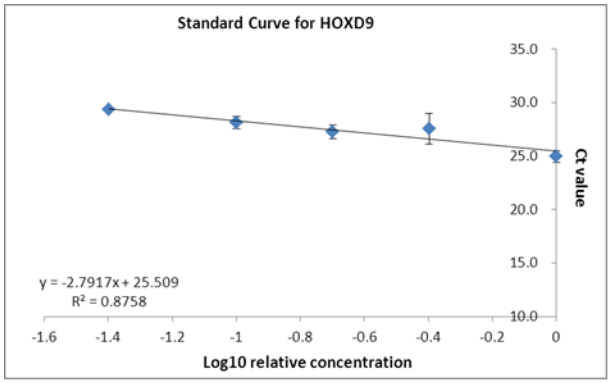
A) Standard Curves for endogenous controls

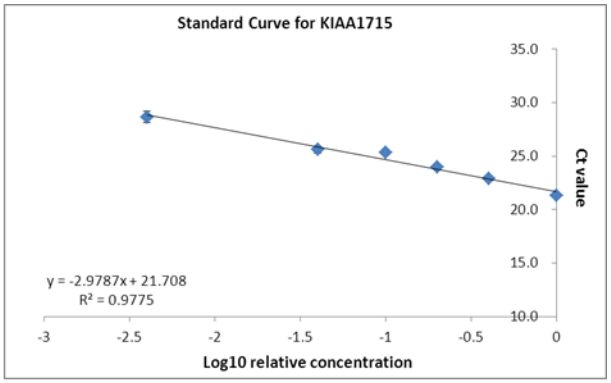


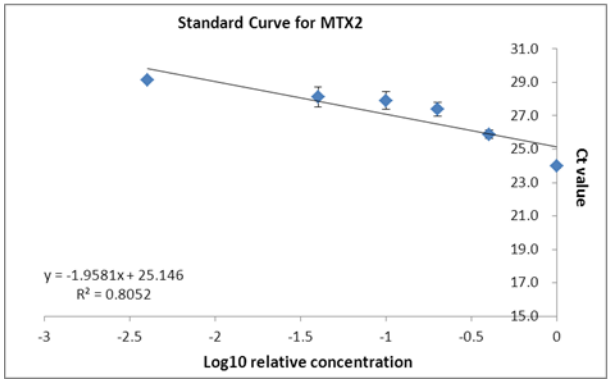
B) Candidate Genes from GWAS Standard Curves for all candidate genes from the 6 loci

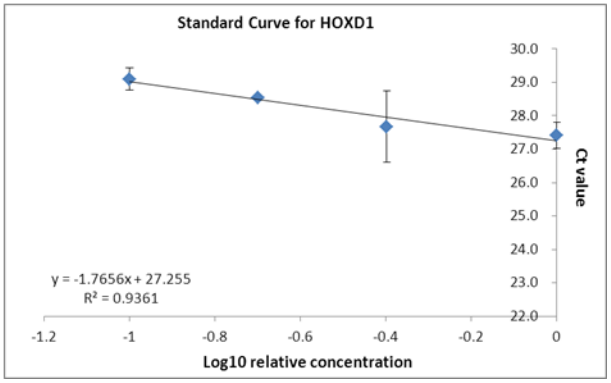
- Chromosome 2



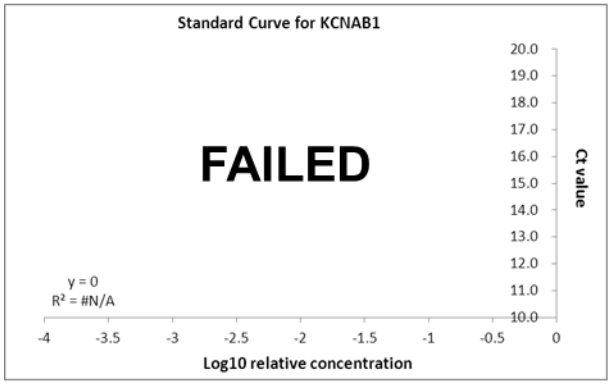


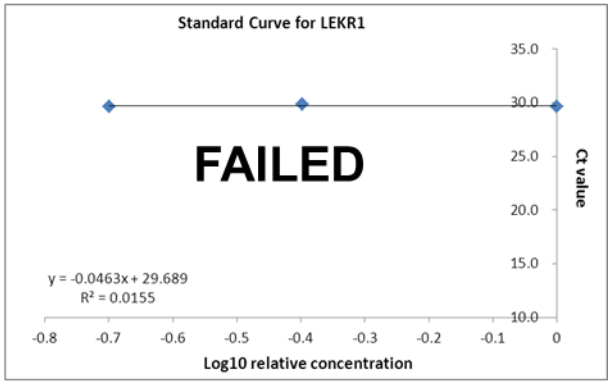


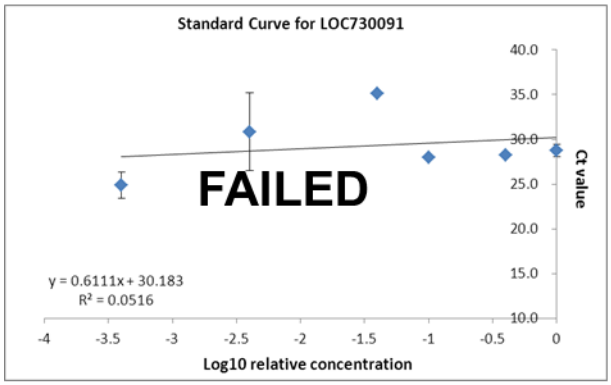


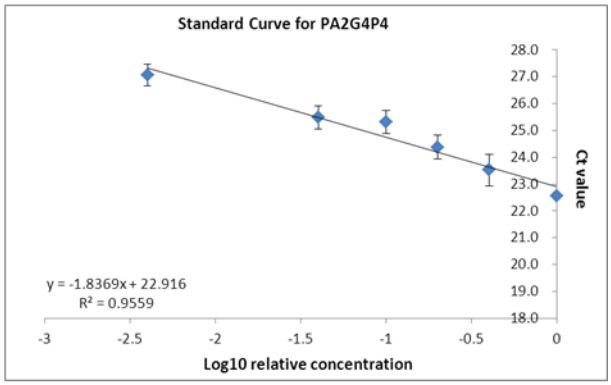


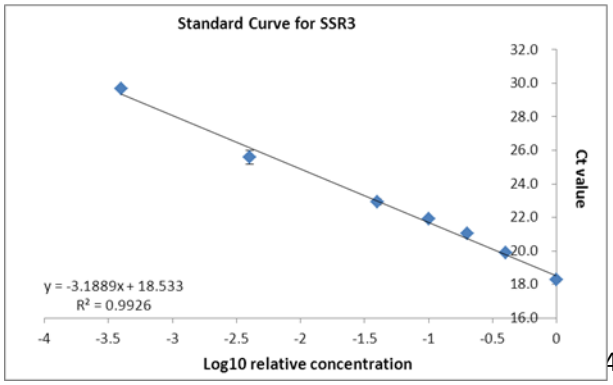
- Chromosome 3

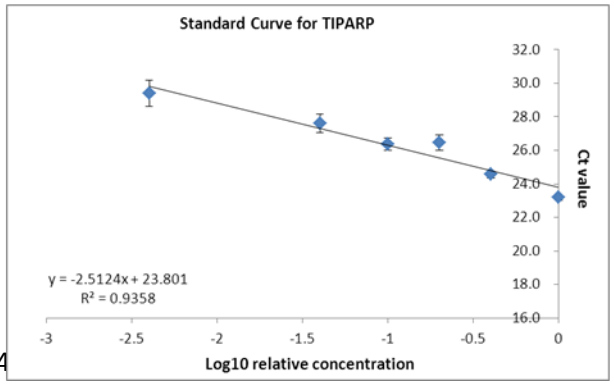




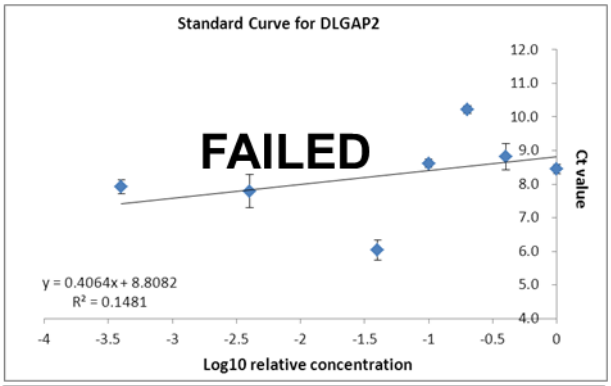


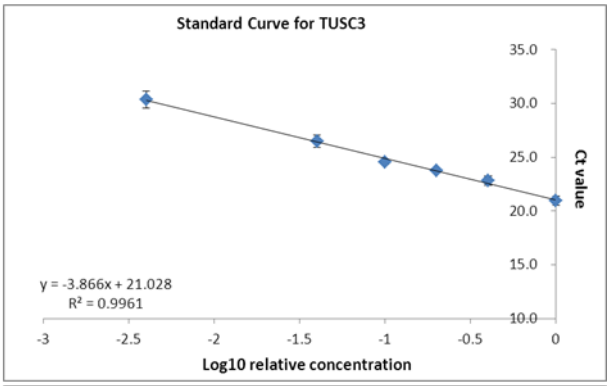


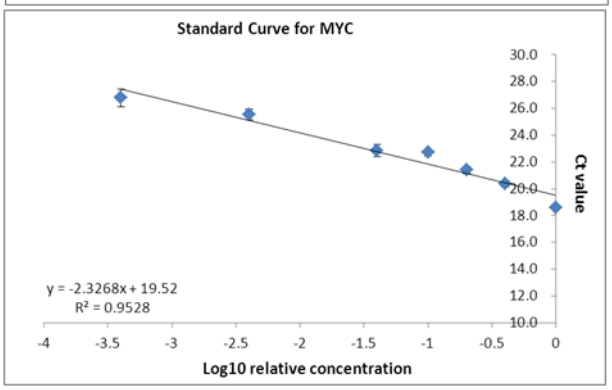


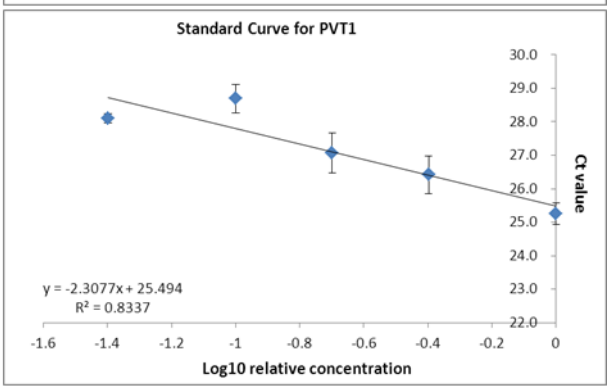


- Chromosome 8

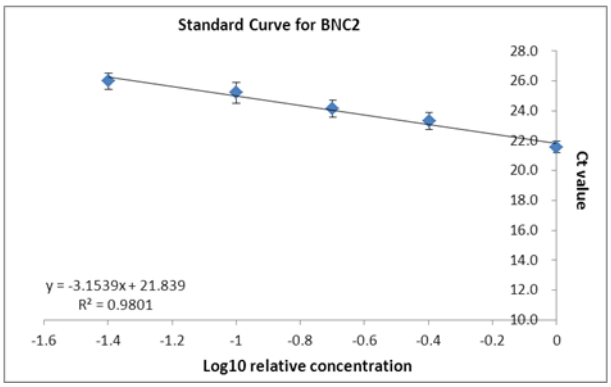


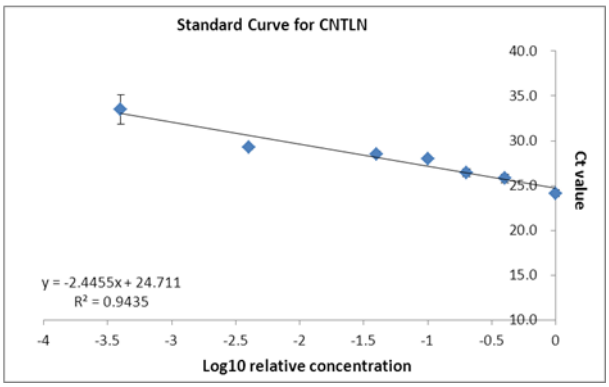




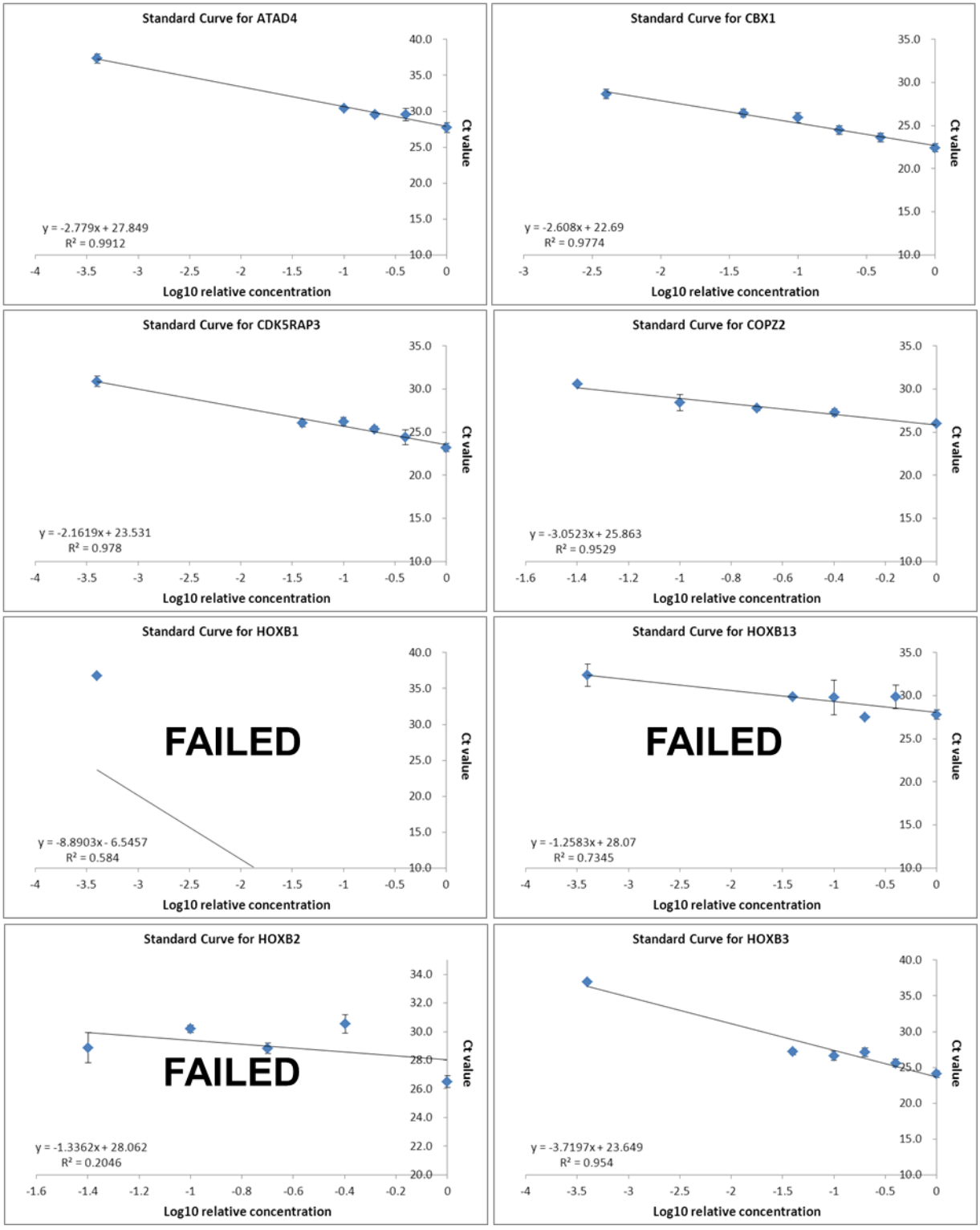


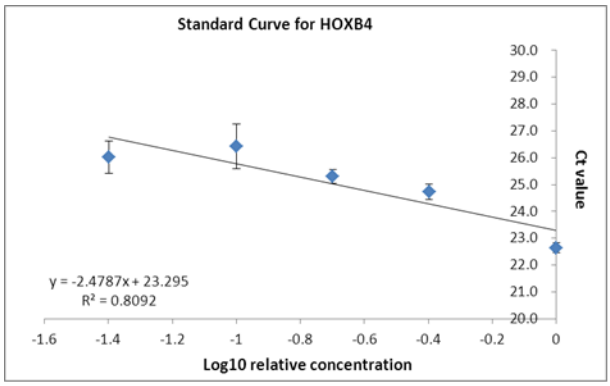
- Chromosome 9

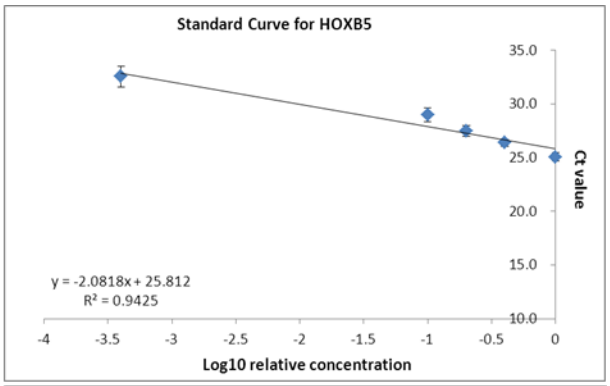


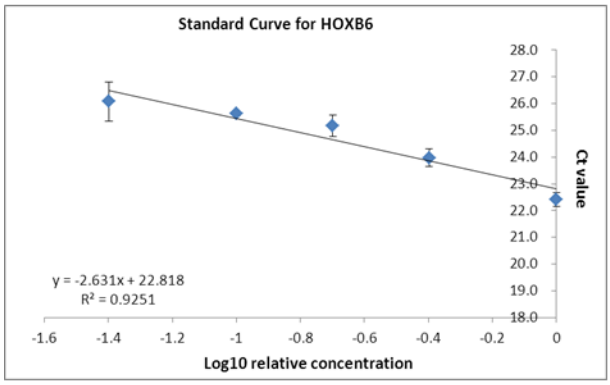


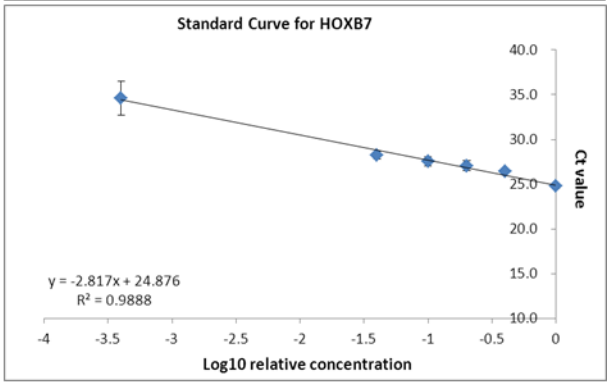
- Chromosome 17

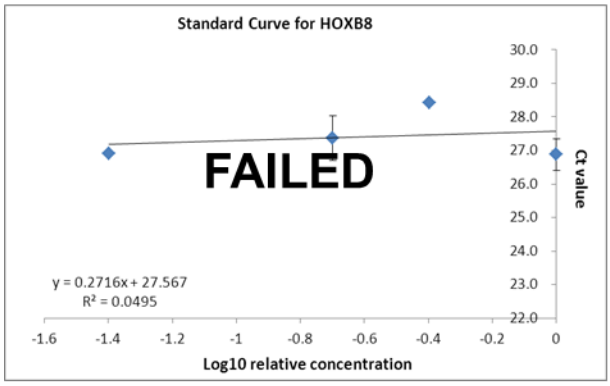


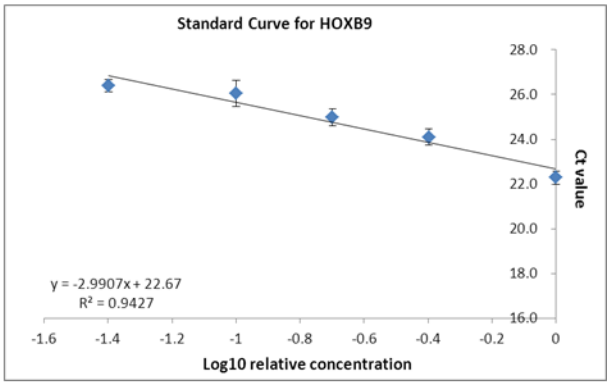

Standard Curve for HOXB4
y = -2.4787x + 23.295
R² = 0.8092
Ct value
Log10 relative concentration

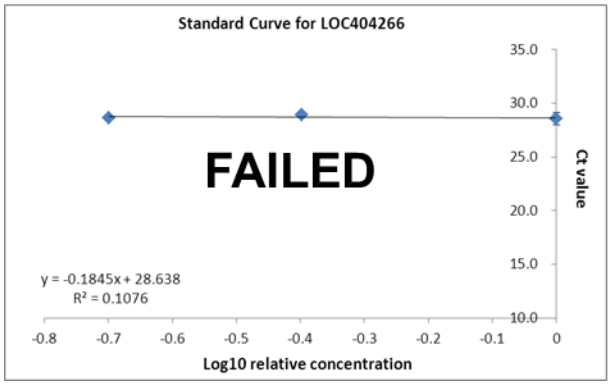

Standard Curve for HOXB5
y = -2.0818x + 25.812
R² = 0.9425
Ct value
Log10 relative concentration

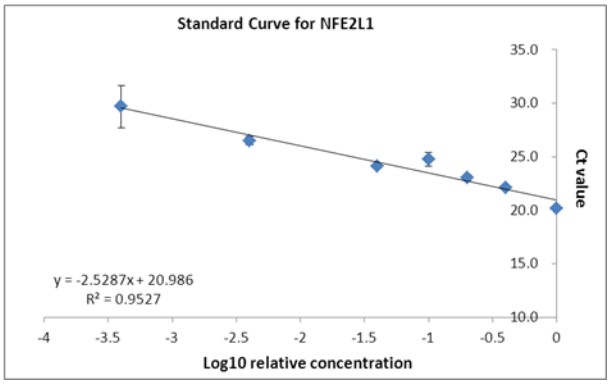

Standard Curve for HOXB6
y = -2.631x + 22.818
R² = 0.9251
Ct value
Log10 relative concentration

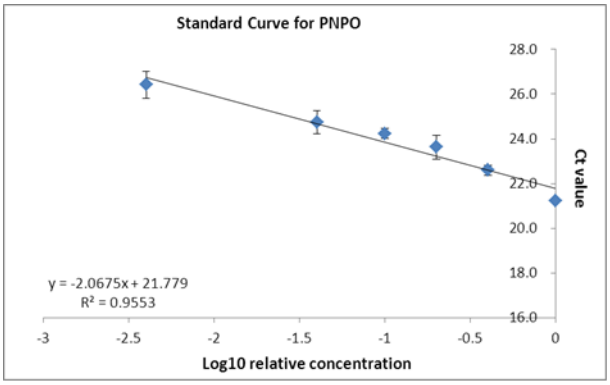

Standard Curve for HOXB7
y = -2.817x + 24.876
R² = 0.9888
Ct value
Log10 relative concentration

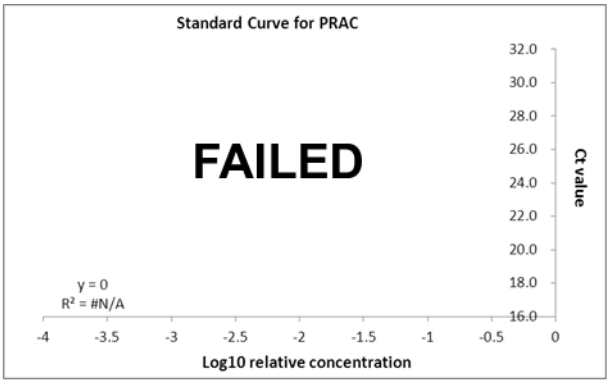

Standard Curve for HOXB8
FAILED
y = 0.2716x + 27.567
R² = 0.0495
Ct value
Log10 relative concentration

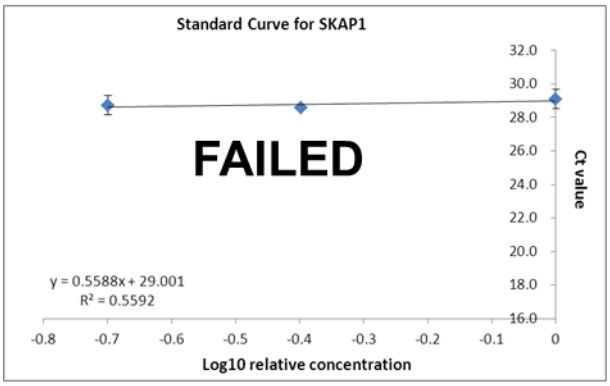

Standard Curve for HOXB9
y = -2.9907x + 22.67
R² = 0.9427
Ct value
Log10 relative concentration

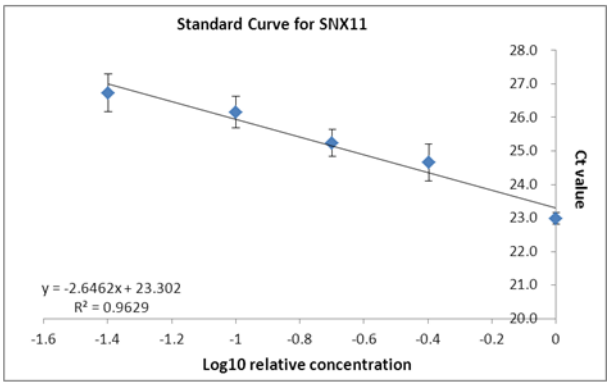

Standard Curve for LOC404266
FAILED
y = -0.1845x + 28.638
R² = 0.1076
Ct value
Log10 relative concentration

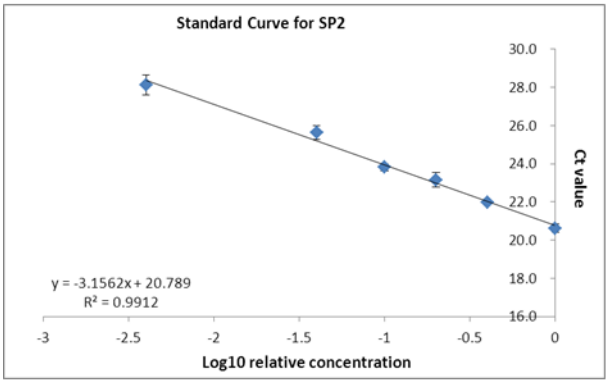

Standard Curve for NFE2L1
y = -2.5287x + 20.986
R² = 0.9527
Ct value
Log10 relative concentration

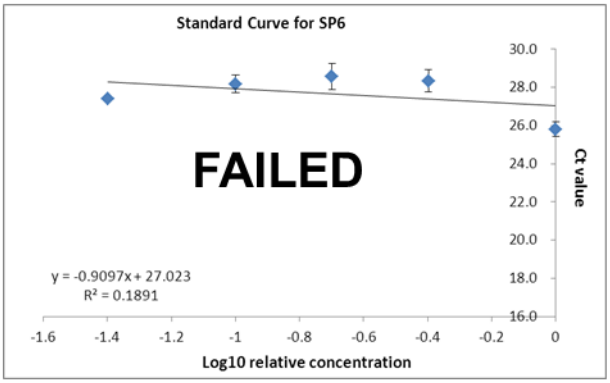

Standard Curve for PNPO
y = -2.0675x + 21.779
R² = 0.9553
Log10 relative concentration
Ct value

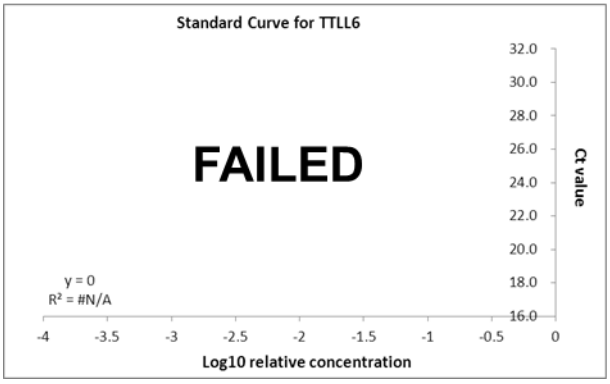

Standard Curve for PRAC
FAILED
y = 0
R² = #N/A
Log10 relative concentration
Ct value


Standard Curve for SKAP1
FAILED
y = 0.5588x + 29.001
R² = 0.5592
Log10 relative concentration
Ct value

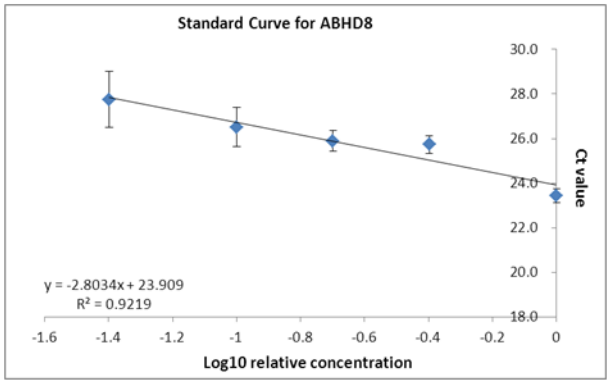

Standard Curve for SNX11
y = -2.6462x + 23.302
R² = 0.9629
Log10 relative concentration
Ct value

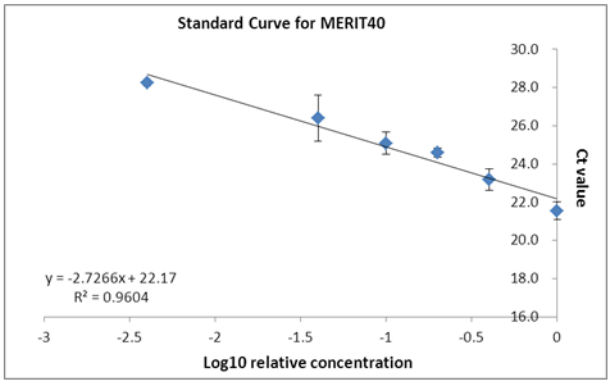

Standard Curve for SP2
y = -3.1562x + 20.789
R² = 0.9912
Log10 relative concentration
Ct value

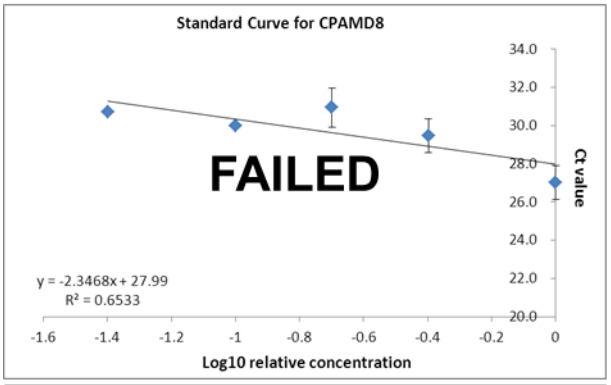

Standard Curve for SP6
FAILED
y = -0.9097x + 27.023
R² = 0.1891
Log10 relative concentration
Ct value

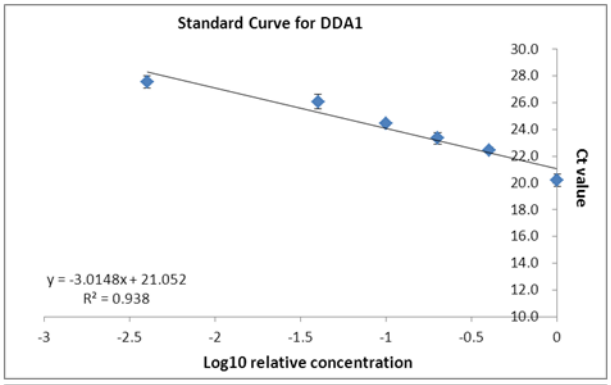

Standard Curve for TTLL6
FAILED
y = 0
R² = #N/A
Log10 relative concentration
Ct value

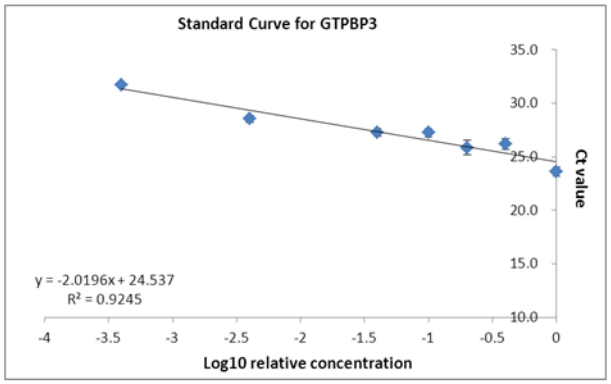
- Chromosome 19

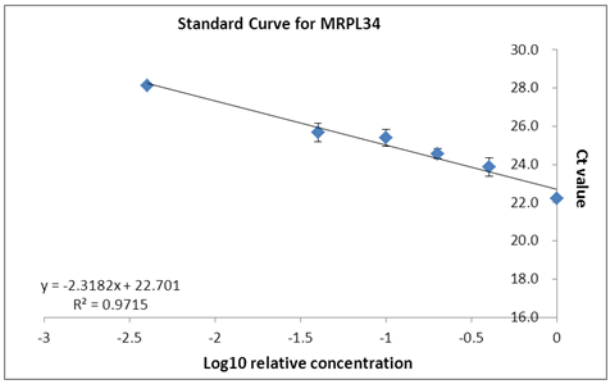


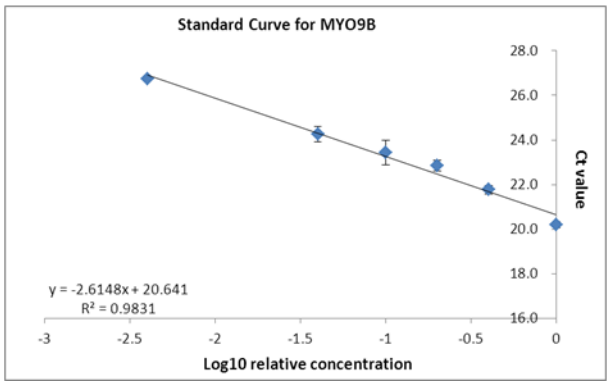


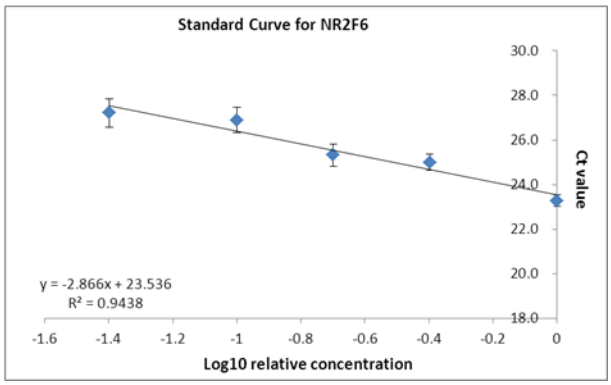


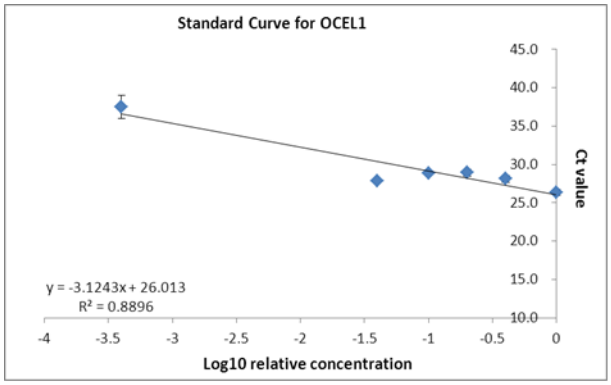


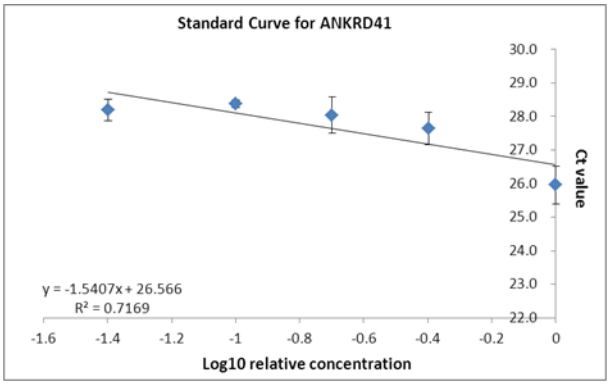


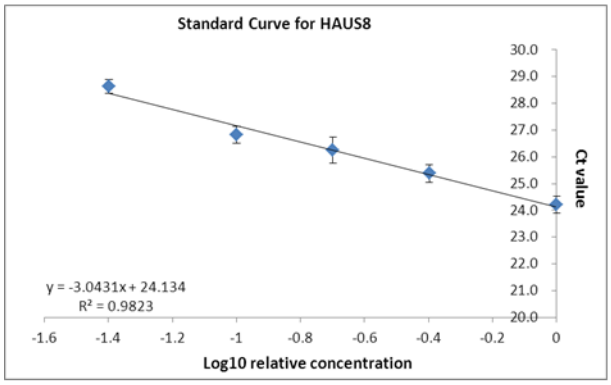


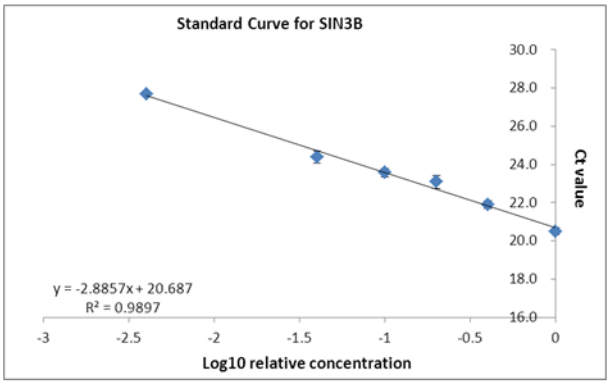


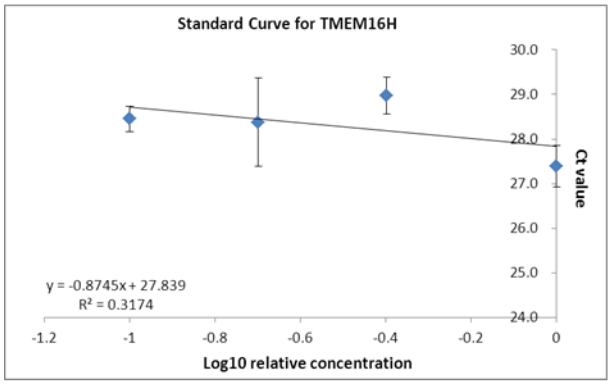


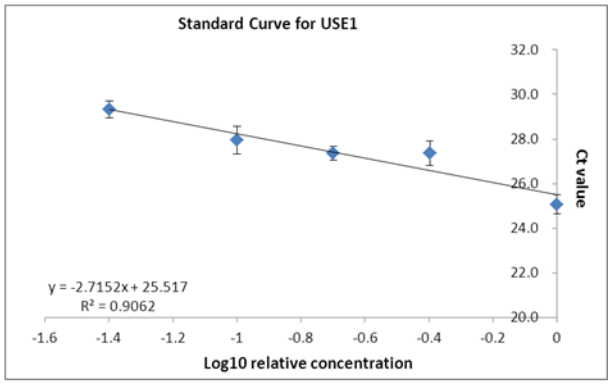

Standard Curve for OCEL1
y = -3.1243x + 26.013
R² = 0.8896
Ct value
Log10 relative concentration


Standard Curve for ANKRD41
y = -1.5407x + 26.566
R² = 0.7169
Ct value
Log10 relative concentration


Standard Curve for HAUS8
y = -3.0431x + 24.134
R² = 0.9823
Ct value
Log10 relative concentration


Standard Curve for SIN3B
y = -2.8857x + 20.687
R² = 0.9897
Ct value
Log10 relative concentration


Standard Curve for TMEM16H
y = -0.8745x + 27.839
R² = 0.3174
Ct value
Log10 relative concentration


Standard Curve for USE1
y = -2.7152x + 25.517
R² = 0.9062
Ct value
Log10 relative concentration

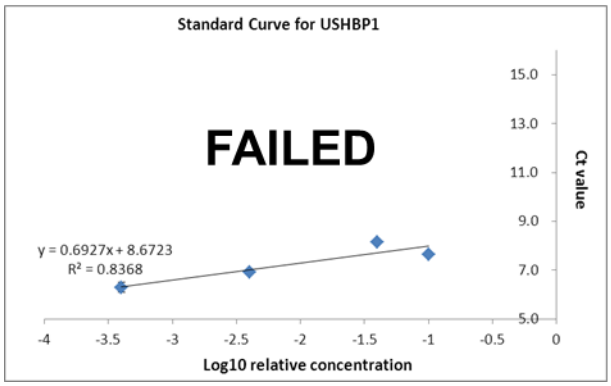

Standard Curve for USHBP1
FAILED
y = 0.6927x + 8.6723
R² = 0.8368
Ct value
Log10 relative concentration

Table 1: Summary of the P values generated for expression of the candidate genes normalised to β-actin and GAPDH separately and combined. Included in this table are genes that failed QC for the Fluidigm assay and their expression between SOC and normal FT from TCGA data as well as reported somatic alterations.

			Fluidigm gene expression assay						TCGA data epression analysis	TCGA analysis for somatic alterations	
			Normalised against β-actin		Normalised against GAPDH		Normalised against β-actin & GAPDH				
Chromosome	Gene	Locus	Sample number (EOC/ Normal)	P value	Sample number (EOC/ Normal)	P value	Sample number (EOC/ Normal)	P value	P value	Number of samples with alterations * (Amp/ HomDel/ Mutations)	Frequency of alterations **
2	KIAA1715	2q31	44/62	0.094	42/58	0.038	42/59	0.004	1.89×10^{-4}	11/0/0	3.5%
	EVX2	2q31.1	FAILED	FAILED	FAILED	FAILED	FAILED	FAILED	0.003	11/0/1	3.8%
	HOXD13	2q31.1	FAILED	FAILED	FAILED	FAILED	FAILED	FAILED	0.007	12/0/0	3.8%
	HOXD12	2q31.1	FAILED	FAILED	FAILED	FAILED	FAILED	FAILED	0.151	12/0/1	4.1%
	HOXD11	2q31.1	FAILED	FAILED	FAILED	FAILED	FAILED	FAILED	0.378	12/0/0	3.8%
	HOXD10	2q31.1	FAILED	FAILED	FAILED	FAILED	FAILED	FAILED	0.024	12/0/3	4.7%
	HOXD9	2q31.1	FAILED	FAILED	FAILED	FAILED	FAILED	FAILED	0.489	12/0/0	3.8%
	HOXD8	2q31.1	30/61	0.185	28/56	0.692	29/59	0.591	0.245	14/0/0	4.4%
	HOXD4	2q31.1	FAILED	FAILED	FAILED	FAILED	FAILED	FAILED	0.663	13/0/0	4.1%
	HOXD3	2q31.1	FAILED	FAILED	FAILED	FAILED	FAILED	FAILED	0.362	14/0/0	4.4%
	HOXD1	2q31.1	FAILED	FAILED	FAILED	FAILED	FAILED	FAILED	0.912	13/0/0	4.1%
	MTX2	2q31.1	44/62	0.001	42/58	0.083	42/59	0.015	0.004	13/0/1	4.4%
3	KCNAB1	3q26.1	FAILED	FAILED	FAILED	FAILED	FAILED	FAILED	2.38×10^{-4}	22/0/2	7.6%
	SSR3	3q25.31	44/62	0.180	42/58	3.04×10^{-4}	42/58	0.003	0.216	22/0/0	7.0%
	TIPARP	3q25.31	41/62	9.4×10^{-6}	42/58	6.57×10^{-11}	41/58	4.99×10^{-9}	1.11×10^{-4}	22/0/1	7.3%
	LOC730091	3q25.31	FAILED	FAILED	FAILED	FAILED	FAILED	FAILED	No data	0/0/0	0.0%
	PA2G4P4	3q25.31	43/62	8.5×10^{-8}	42/58	3.64×10^{-5}	41/61	2.14×10^{-5}	No data	22/0/1	7.3%
	LEKR1	3q25.31	FAILED	FAILED	FAILED	FAILED	FAILED	FAILED	0.006	23/0/1	7.6%
8	MYC	8q24.21	44/62	0.028	39/57	0.079	40/59	0.085	0.059	97/0/0	30.7%
	PVT1	8q24	42/62	9.7×10^{-11}	41/56	3.28×10^{-8}	41/59	6.64×10^{-8}	No data	102/0/0	32.3%
9	BNC2	9p22.32	35/62	N/A *	30/58	1.73×10^{-4}	32/58	N/A *	1.63×10^{-4}	8/2/1	3.5%
	CNTLN	9p22	44/62	N/A *	42/58	0.030	42/59	N/A *	0.536	4/3/0	2.2%
17	SP6	17q21.32	FAILED	FAILED	FAILED	FAILED	FAILED	FAILED	0.005	7/0/0	2.2%
	SP2	17q21.32	44/62	1.69×10^{-12}	42/57	2.33×10^{-6}	42/58	5.68×10^{-10}	0.049	6/0/0	1.9%
	PNPO	17q21.32	44/62	8.59×10^{-10}	42/57	8.21×10^{-5}	42/58	6.26×10^{-8}	0.629	6/0/0	1.9%
	ATAD4	17q21.32	FAILED	FAILED	FAILED	FAILED	FAILED	FAILED	0.210	6/0/1	2.2%
	CDK5RAP3	17q21.32	44/62	N/A *	42/57	0.023	42/51	N/A *	0.143	5/0/0	1.6%
	COPZ2	17q21.32	30/62	N/A *	29/57	0.001	29/58	N/A *	0.107	5/0/0	1.6%
	NFE2L1	17q21.3	44/62	0.465	42/58	0.088	41/59	0.050	1.4×10^{-5}	5/0/0	1.6%
	CBX1	17q21.32	44/62	N/A *	42/57	1.19×10^{-5}	42/58	N/A *	0.018	6/0/0	1.9%
	SNX11	17q21.32	43/62	3.77×10^{-7}	42/58	0.532	42/58	0.004	0.388	5/0/0	1.6%
	SKAP1	17q21.32	FAILED	FAILED	FAILED	FAILED	FAILED	FAILED	0.001	10/1/0	3.5%
	HOXB1	17q21.3	FAILED	FAILED	FAILED	FAILED	FAILED	FAILED	0.024	10/0/1	3.5%
	HOXB2	17q21-q22	FAILED	FAILED	FAILED	FAILED	FAILED	FAILED	0.047	9/0/0	2.8%
	HOXB3	17q21.3	34/52	0.002	28/47	1.54×10^{-5}	31/51	1.72×10^{-4}	5.24×10^{-4}	10/0/0	3.2%
	HOXB4	17q21-q22	34/54	6.8×10^{-7}	33/49	4.9×10^{-4}	35/52	0.001	0.026	10/0/1	3.5%
	HOXB5	17q21.3	FAILED	FAILED	FAILED	FAILED	FAILED	FAILED	0.002	9/0/0	2.8%
	HOXB6	17q21.3	27/54	1.31×10^{-4}	21/49	4.76×10^{-5}	26/50	4.23×10^{-7}	5.25×10^{-5}	8/0/0	2.5%
	LOC404266	17q21.32	FAILED	FAILED	FAILED	FAILED	FAILED	FAILED	No data	0/0/0	0.0%
	HOXB7	17q21.3	35/54	5.12×10^{-11}	35/52	4.08×10^{-9}	35/50	4.62×10^{-10}	0.867	8/0/0	2.5%
	HOXB8	17q21.3	FAILED	FAILED	FAILED	FAILED	FAILED	FAILED	0.001	8/0/0	2.5%
	HOXB9	17q21.3	29/47	3.84×10^{-10}	27/46	1.05×10^{-8}	30/46	9.7×10^{-9}	0.946	8/0/0	2.5%
	PRAC	17q21	FAILED	FAILED	FAILED	FAILED	FAILED	FAILED	0.184	7/0/0	2.2%
	HOXB13	17q21.2	FAILED	FAILED	FAILED	FAILED	FAILED	FAILED	0.003	7/0/0	2.2%
	TTLL6	17q21.32	FAILED	FAILED	FAILED	FAILED	FAILED	FAILED	2.12×10^{-6}	7/0/0	2.2%
19	SIN3B	19p13.11	44/62	8.57×10^{-5}	42/57	0.864	42/57	0.126	0.019	28/0/2	9.5%
	CPAMD8	19p13.11	FAILED	FAILED	FAILED	FAILED	FAILED	FAILED	0.398	28/0/6	10.8%
	HAUS8	19p13.11	43/62	2.7×10^{-11}	42/51/6	1.02×10^{-4}	41/62	1.08×10^{-6}	1.48×10^{-5}	27/0/1	8.9%
	MYO9B	19p13.1	44/62	0.071	42/58	0.082	42/59	0.010	3.69×10^{-5}	30/0/2	10.1%
	USE1	19p13.11	44/62	8.52×10^{-9}	42/58	0.002	43/60	1.69×10^{-5}	0.224	30/0/0	9.5%
	OCEL1	19p13.11	39/62	0.002	39/59	0.037	38/59	0.012	0.009	30/0/0	9.5%
	NR2F6	19p13.1	42/62	0.075	40/58	0.562	42/60	0.226	0.094	30/0/0	9.5%
	USHBP1	19p13	FAILED	FAILED	FAILED	FAILED	FAILED	FAILED	0.014	30/0/1	9.8%
	MERIT40	19p13.11	44/62	N/A *	42/57	0.005	42/58	N/A *	2.09×10^{-5}	29/0/1	9.5%
	ANKRD41	19p13.11	FAILED	FAILED	FAILED	FAILED	FAILED	FAILED	0.234	28/0/0	8.9%
	ABHD8	19p13.11	44/62	N/A *	42/58	0.028	42/58	N/A *	0.132	27/0/0	8.5%
	DDA1	19p13.11	44/62	N/A *	42/58	0.471	42/60	N/A *	2.04×10^{-4}	27/0/0	8.5%
	MRPL34	19p13.1	43/62	3.34×10^{-5}	42/58	0.003	42/60	0.001	0.005	27/0/0	8.5%
	TMEM16H	19p13.11	FAILED	FAILED	FAILED	FAILED	FAILED	FAILED	7.42×10^{-5}	26/0/3	9.2%
	GTPBP3	19p13.11	FAILED	FAILED	FAILED	FAILED	FAILED	FAILED	1.02×10^{-5}	25/0/0	7.9%

Table 2: Summary of genotype specific methylation of CpG islands associated with the candidate genes within risk loci in chromosomes 2, 3, 8 and 9 (Linear regression analysis). Tabulated are the CpGs' median methylation (represented by β-values) compared between the 3 genotype groups of the risk EOC associated SNPs. P values were generated with linear regression analysis.

Chromosome	SNP	Genotype (Major/Minor)	Gene	CpG site	Median β-value comparison between common homozygotes and combined rare homozygotes & heterozygotes							
					Sample number CH/H/RH	Median of β-values			Median Absolute Deviation (MAD)			Linear regression analysis (P value)
						CH	H	RH	CH	H	RH	
2	rs2072590	G/T	HOXD9	cg10957151	128/111/17	0.174	0.157	0.157	0.020	0.022	0.017	0.003
				cg14991487		0.053	0.050	0.049	0.007	0.009	0.009	0.332
			HOXD8	cg15520279		0.036	0.033	0.033	0.009	0.008	0.008	0.455
				cg21815667		0.074	0.075	0.074	0.009	0.007	0.007	0.901
			HOXD3	cg00005847		0.150	0.150	0.153	0.017	0.016	0.013	0.186
				cg18702197		0.053	0.054	0.047	0.007	0.007	0.008	0.621
			HOXD1	cg19001226		0.041	0.043	0.036	0.006	0.006	0.005	0.616
			MTX2	cg19280014		0.083	0.084	0.086	0.008	0.009	0.006	0.447
3	rs2665390	T/C	KCNAB1	cg15423862	210/44/2	0.855	0.855	0.883	0.010	0.011	0.009	0.083
				cg23873703		0.695	0.703	0.721	0.058	0.062	0.069	0.489
			SSR3	cg24517609		0.073	0.071	0.072	0.006	0.006	0.010	0.977
			TIPARP	cg22114222		0.022	0.021	0.021	0.002	0.002	0.001	0.389
				cg24262469		0.196	0.208	0.205	0.022	0.025	0.015	0.304
8	rs10088218	G/A	MYC	cg19972619	198/50/8	0.044	0.044	0.048	0.004	0.003	0.003	0.817
				cg27207274		0.071	0.072	0.072	0.006	0.006	0.005	0.222
			PVT1	cg13052755		0.886	0.887	0.884	0.014	0.011	0.008	0.415
				cg13784855		0.038	0.037	0.038	0.003	0.003	0.002	0.993
9	rs3814113	T/C	BNC2	cg24341129	108/117/31	0.276	0.282		0.035	0.032		0.366

Table 3: Summary of genotype specific methylation of CpG islands associated with the candidate genes within loci in chromosome 17 (Linear regression analysis). Tabulated are the CpGs' median methylation (represented by β-values) compared between the 3 genotype groups of the risk EOC associated SNPs. P values were generated with linear regression analysis.

Chromosome	SNP	Genotype (Major/Minor)	Gene	CpG site	Median β-value comparison between common homozygotes and combined rare homozygotes & heterozygotes							
					Sample number CH/H/RH	Median of β-values			Median Absolute Deviation (MAD)			Linear regression analysis
						CH	H	RH	CH	H	RH	
17	rs9303542	A/G	SP2	cg02691091	129/105/22	0.906	0.904	0.907	0.007	0.008	0.008	0.646
				cg14360917		0.530	0.517	0.532	0.040	0.027	0.030	0.151
			PNPO	cg00177698		0.037	0.035	0.037	0.004	0.004	0.002	0.760
				cg10154655		0.041	0.042	0.041	0.004	0.003	0.003	0.185
			NFE2L1	cg03777405		0.062	0.064	0.066	0.008	0.007	0.013	0.262
				cg06740897		0.083	0.086	0.080	0.012	0.015	0.014	0.665
			CBX1	cg11194725		0.053	0.052	0.052	0.005	0.005	0.004	0.271
				cg17778721		0.040	0.040	0.040	0.006	0.006	0.007	0.719
			SNX11	cg08413153		0.111	0.116	0.116	0.012	0.013	0.010	0.340
				cg12563480		0.035	0.036	0.035	0.006	0.006	0.005	0.601
			SKAP1	cg05106502		0.368	0.380	0.374	0.039	0.033	0.034	0.763
				cg12513481		0.067	0.067	0.066	0.012	0.008	0.014	0.917
			HOXB2	cg25882366		0.260	0.272	0.278	0.044	0.052	0.037	0.604
			HOXB3	cg12910797		0.841	0.847	0.849	0.021	0.016	0.021	0.998
			HOXB4	cg02422694		0.040	0.040	0.040	0.004	0.004	0.003	0.798
				cg04609859		0.034	0.033	0.034	0.005	0.004	0.005	0.677
				cg06760035		0.033	0.034	0.036	0.003	0.003	0.005	0.005
				cg08089301		0.035	0.036	0.034	0.004	0.005	0.006	0.496
				cg14458834		0.118	0.127	0.126	0.019	0.021	0.028	0.320
				cg21460081		0.078	0.080	0.081	0.007	0.008	0.009	0.437
				cg21546671		0.115	0.120	0.115	0.016	0.022	0.020	0.969
				cg25145670		0.075	0.076	0.091	0.010	0.009	0.013	0.013
			HOXB5	cg01405107		0.273	0.239	0.219	0.041	0.051	0.040	5.56×10^{-6}
				cg16495265		0.081	0.082	0.079	0.008	0.008	0.008	0.451
			HOXB6	cg16848873		0.456	0.493	0.562	0.080	0.072	0.079	0.006
			HOXB7	cg06493080		0.051	0.051	0.048	0.008	0.010	0.007	0.948
				cg09357097		0.062	0.062	0.062	0.007	0.007	0.008	0.772
			HOXB9	cg12370791		0.030	0.031	0.029	0.003	0.003	0.003	0.972
				cg13643585		0.028	0.027	0.026	0.003	0.003	0.004	0.335

Table 4: Summary of genotype specific methylation of CpG islands associated with the candidate genes within loci in chromosome 19 (Linear regression analysis). Tabulated are the CpGs' median methylation (represented by β-values) compared between the 3 genotype groups of the risk EOC associated SNPs. P values were generated with linear regression analysis.

Chromosome	SNP	Genotype (Major/Minor)	Gene	CpG site	Median β-value comparison between common homozygotes and combined rare homozygotes & heterozygotes							
					Sample number CH/H/RH	Median of β-values			Median Absolute Deviation (MAD)			Linear regression analysis (P value)
						CH	H	RH	CH	H	RH	
19	rs2363956	T/G	MYO9B	cg00215066	56/123/77	0.949	0.950	0.950	0.005	0.005	0.005	0.728
				cg01103836		0.962	0.963	0.962	0.003	0.003	0.004	0.195
			USE1	cg01252496		0.510	0.505	0.500	0.015	0.019	0.020	0.092
				cg08891270		0.035	0.034	0.034	0.002	0.003	0.003	0.584
			NR2F6	cg07242565		0.033	0.033	0.033	0.003	0.003	0.003	0.892
				cg16749578		0.166	0.166	0.160	0.022	0.024	0.020	0.442
			ANKRD41	cg21435336		0.037	0.035	0.035	0.003	0.003	0.003	0.045
			ABHD8	cg08145177		0.117	0.120	0.110	0.019	0.030	0.020	0.666
				cg22734480		0.027	0.025	0.026	0.002	0.003	0.003	0.041
			DDA1	cg19390148		0.032	0.031	0.030	0.005	0.005	0.004	0.425
				cg21061811		0.747	0.740	0.735	0.021	0.022	0.028	0.397
			MRPL34	cg04914305		0.050	0.050	0.050	0.005	0.004	0.004	0.242
	rs8170	T/C	MYO9B	cg00215066	180/69/7	0.949	0.950	0.950	0.005	0.005	0.005	0.728
				cg01103836		0.962	0.963	0.962	0.003	0.003	0.004	0.195
			USE1	cg01252496		0.510	0.505	0.500	0.015	0.019	0.020	0.092
				cg08891270		0.035	0.034	0.034	0.002	0.003	0.003	0.584
			NR2F6	cg07242565		0.033	0.033	0.033	0.003	0.003	0.003	0.892
				cg16749578		0.166	0.166	0.160	0.022	0.024	0.020	0.442
			ANKRD41	cg21435336		0.037	0.035	0.035	0.003	0.003	0.003	0.045
			ABHD8	cg08145177		0.117	0.120	0.110	0.019	0.030	0.020	0.666
				cg22734480		0.027	0.025	0.026	0.002	0.003	0.003	0.041
			DDA1	cg19390148		0.032	0.031	0.030	0.005	0.005	0.004	0.425
				cg21061811		0.747	0.740	0.735	0.021	0.022	0.028	0.397
			MRPL34	cg04914305		0.050	0.050	0.050	0.005	0.004	0.004	0.242

Appendix 4: *Functional analysis of MERIT40*

Figure 1: Cell cycle profiles of A2780CP, MPSC1, OVCAR3 control and *MERIT40* knockdown cell lines

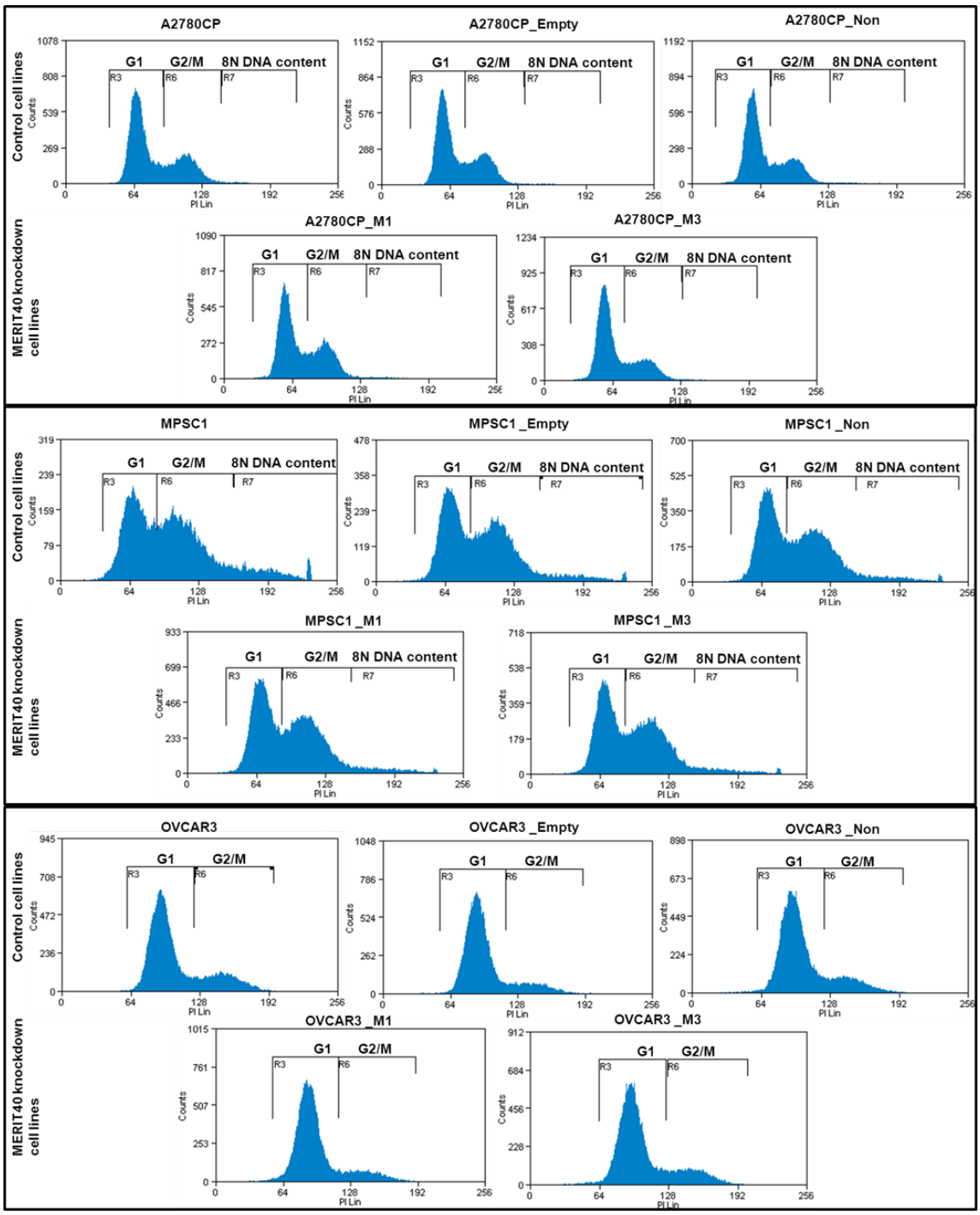


Figure 2: Representative soft agar assays of EFO27, A2780CP and MPSC1 control and *MERIT40* knockdown cell lines. The pictured were taken after seeding 5000 cells and incubating for 4 weeks. 112D EOC cell line was used as a positive control cell line.

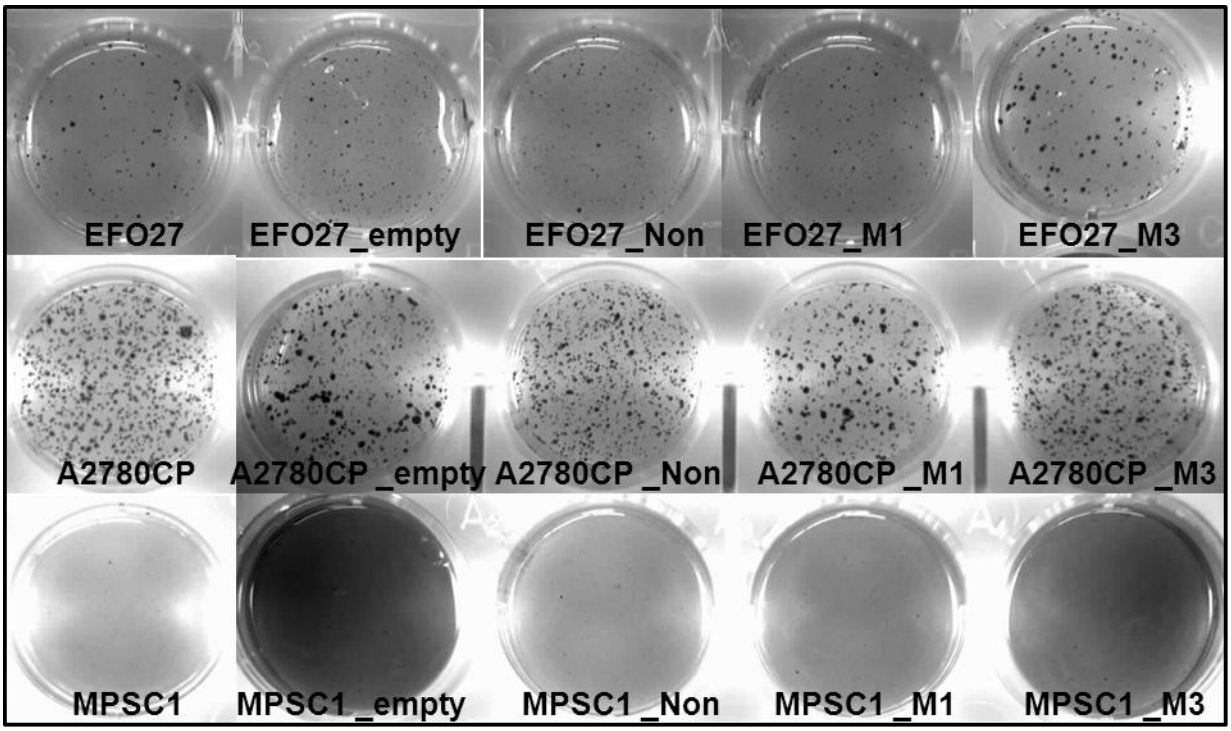
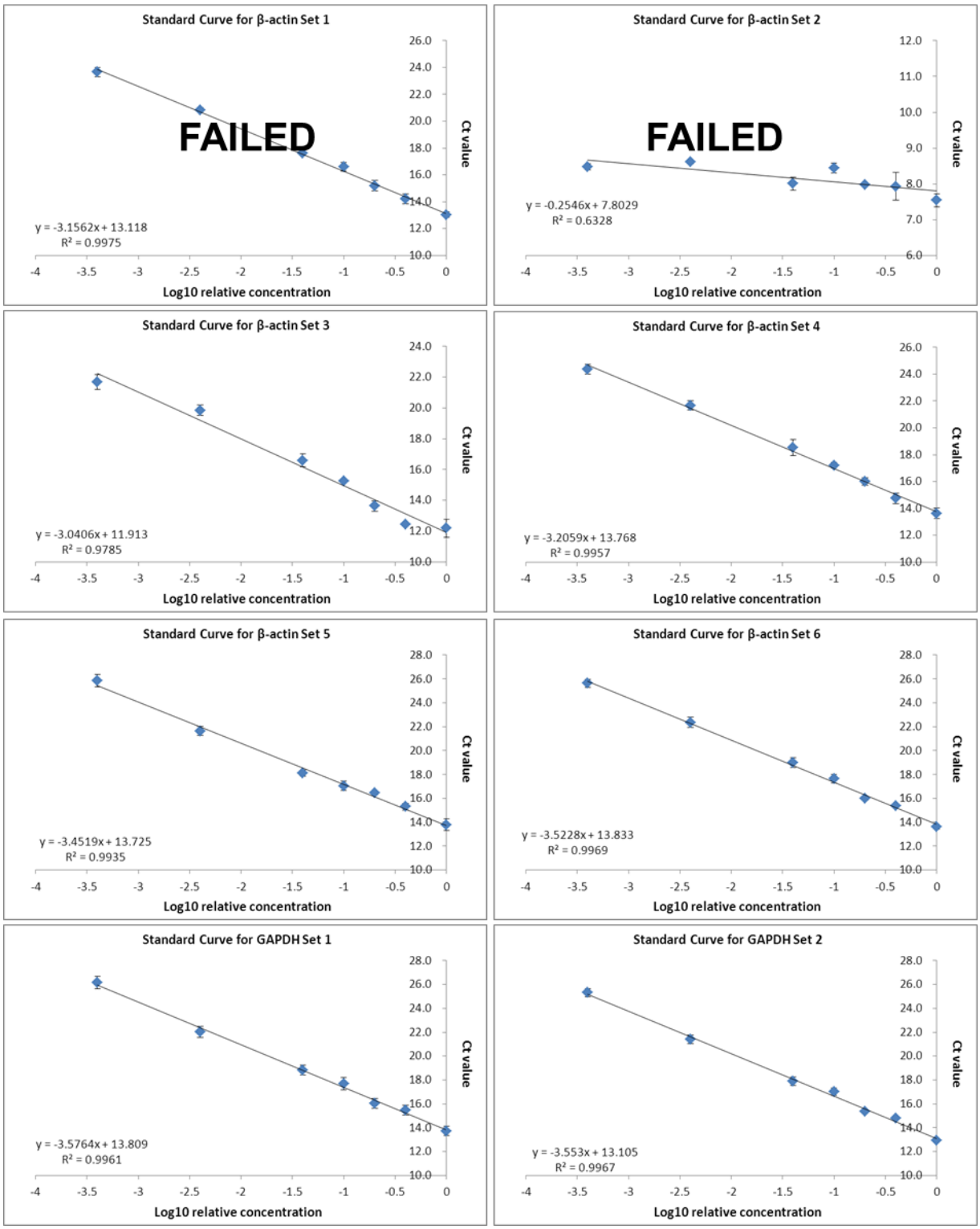
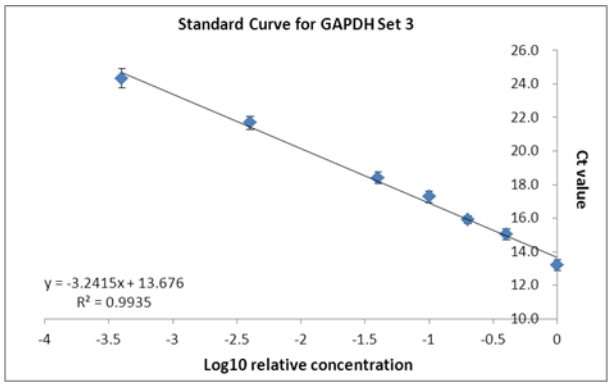
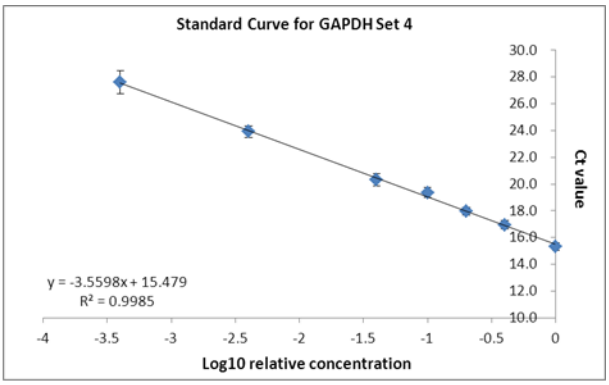


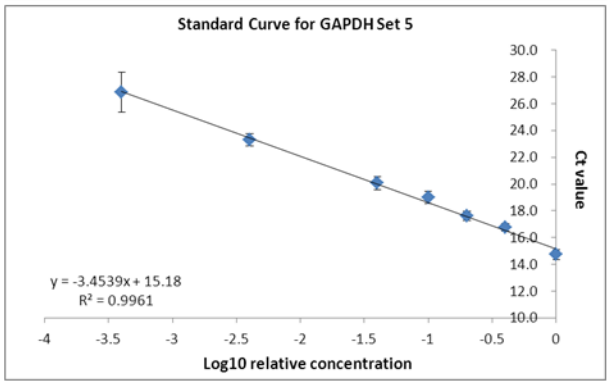
Figure 3: Standard Curves for assays run in the Fluidigm assay for gene expression analysis between EOC control and *MERIT40* knockdown cell lines

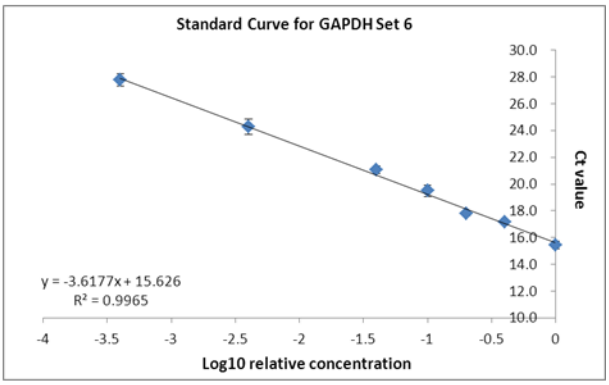
A) Endogenous controls

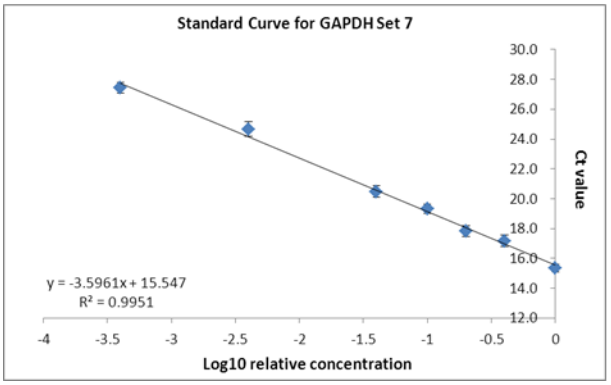



Standard Curve for GAPDH Set 3
y = -3.2415x + 13.676
R² = 0.9935
Ct value
Log10 relative concentration

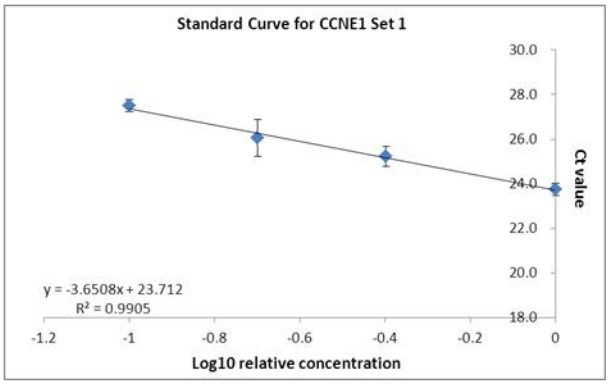

Standard Curve for GAPDH Set 4
y = -3.5598x + 15.479
R² = 0.9985
Ct value
Log10 relative concentration

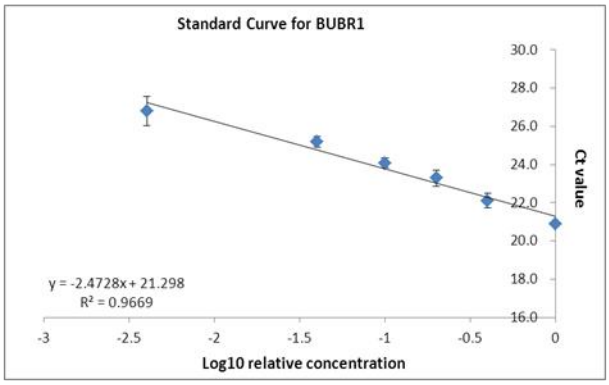

Standard Curve for GAPDH Set 5
y = -3.4539x + 15.18
R² = 0.9961
Ct value
Log10 relative concentration

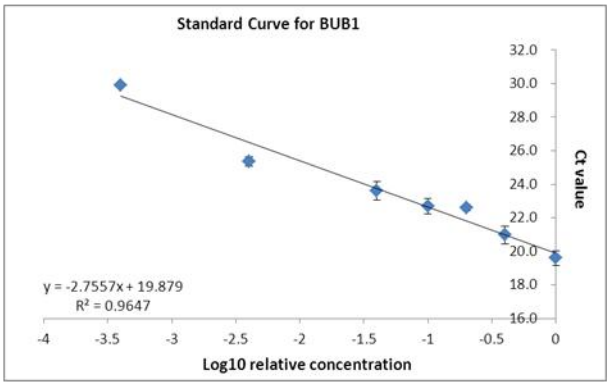

Standard Curve for GAPDH Set 6
y = -3.6177x + 15.626
R² = 0.9965
Ct value
Log10 relative concentration

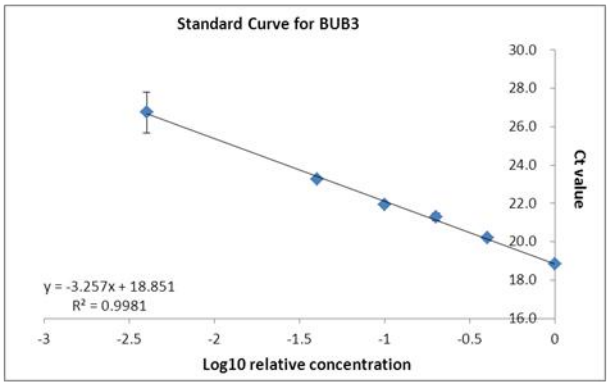

Standard Curve for GAPDH Set 7
y = -3.5961x + 15.547
R² = 0.9951
Ct value
Log10 relative concentration

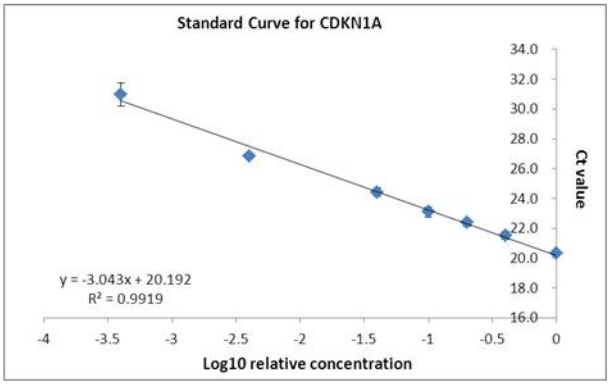
B) Genes involved in cell cycle regulation and chromosomal segregation

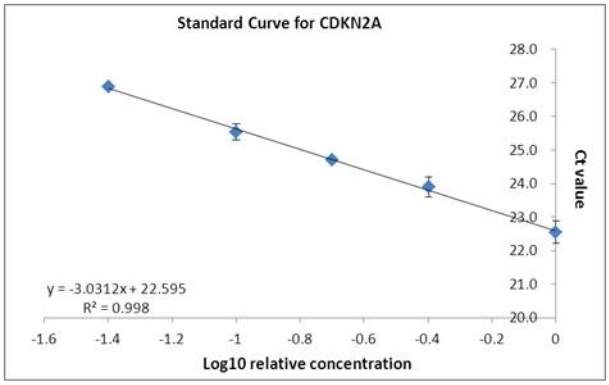


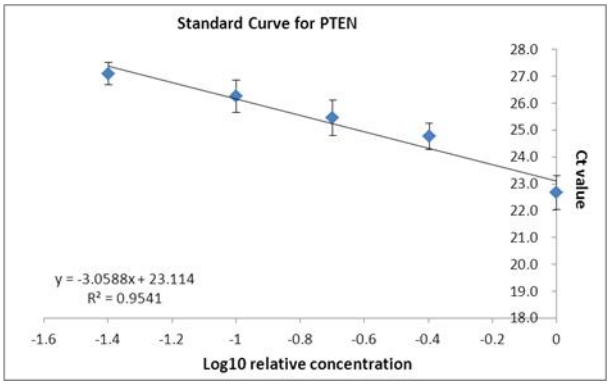


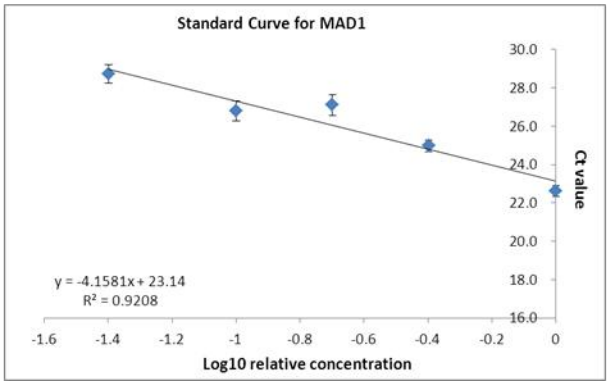


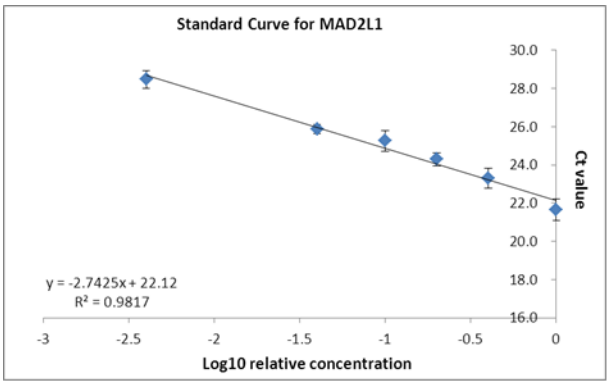


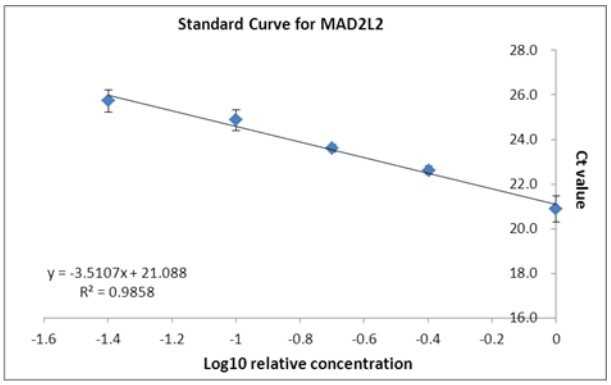


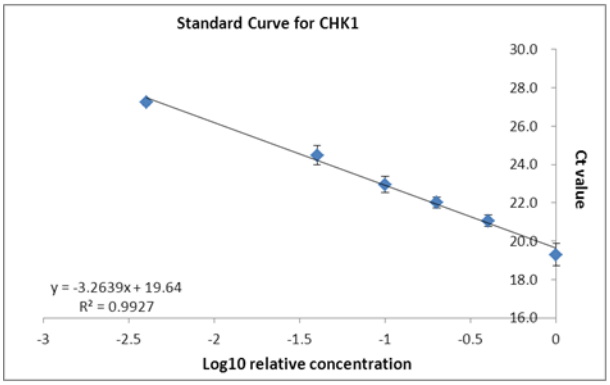


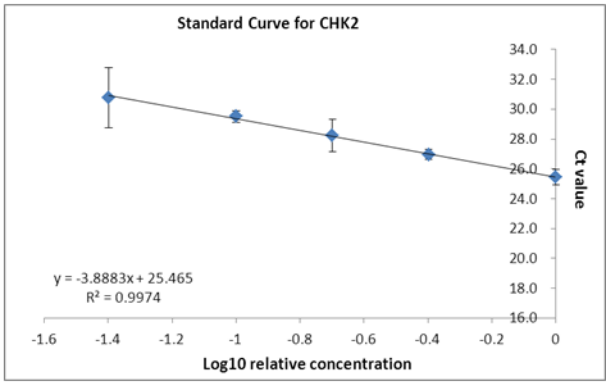


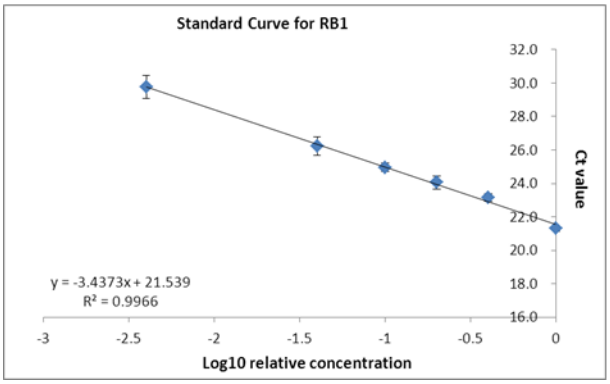



Standard Curve for MAD2L1
y = -2.7425x + 22.12
R² = 0.9817
Ct value
Log10 relative concentration


Standard Curve for MAD2L2
y = -3.5107x + 21.088
R² = 0.9858
Ct value
Log10 relative concentration

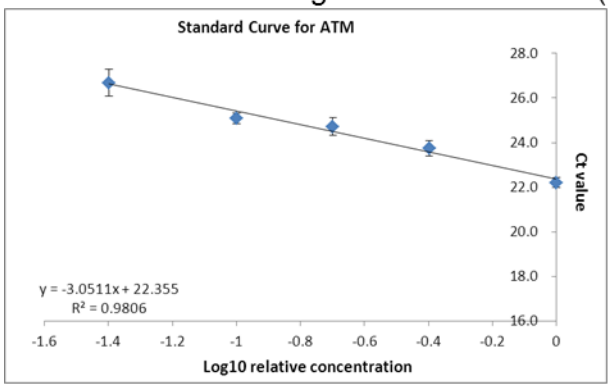

Standard Curve for CHK1
y = -3.2639x + 19.64
R² = 0.9927
Ct value
Log10 relative concentration

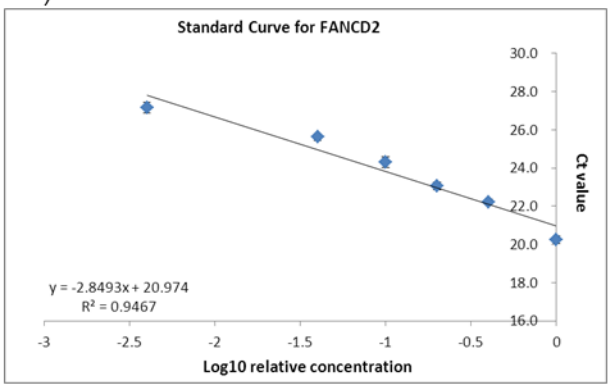

Standard Curve for CHK2
y = -3.8883x + 25.465
R² = 0.9974
Ct value
Log10 relative concentration

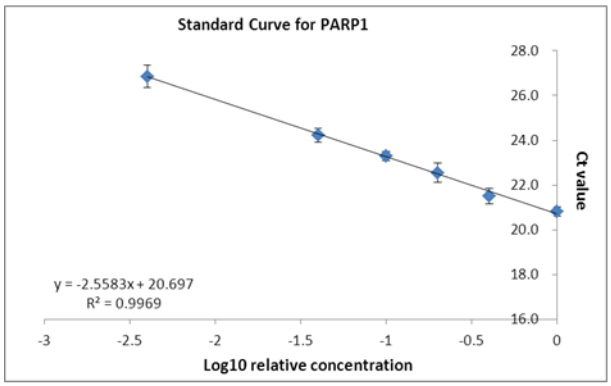

Standard Curve for RB1
y = -3.4373x + 21.539
R² = 0.9966
Ct value
Log10 relative concentration

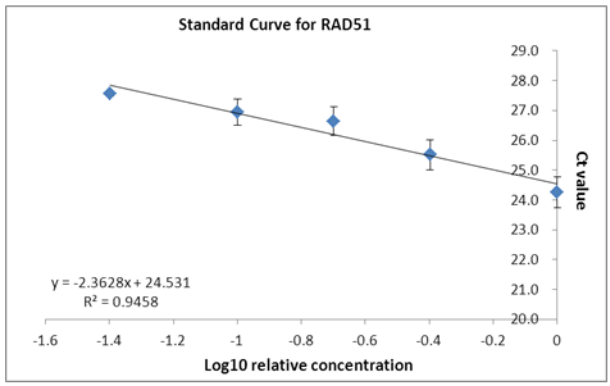
C) Genes involved in DNA damage responses and genes encoding MERIT40 interacting proteins within the BRCA1 complex

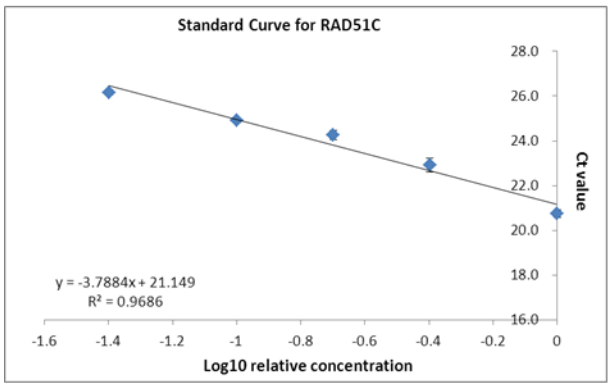
- Involved in homologous recombination (HR)

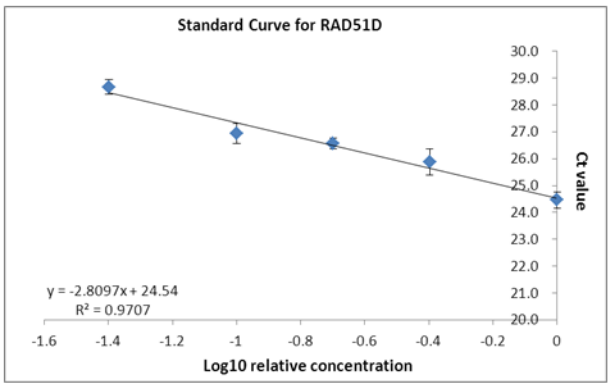


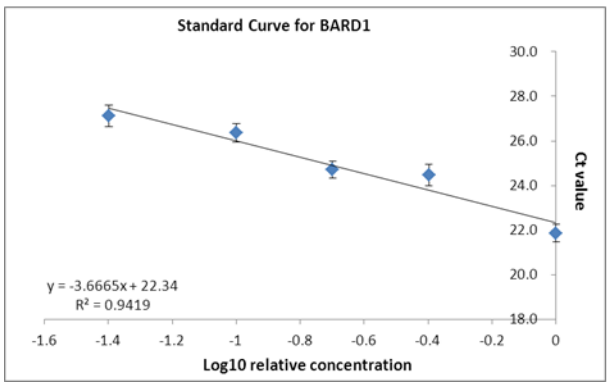


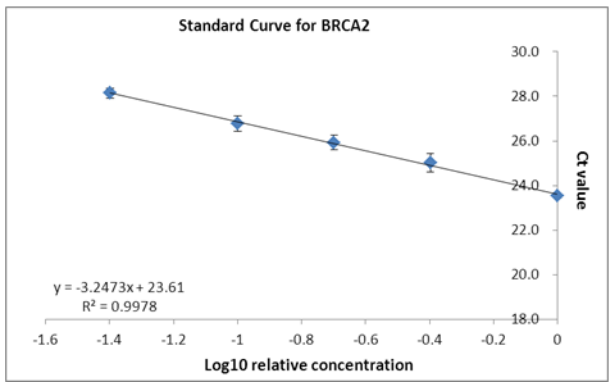






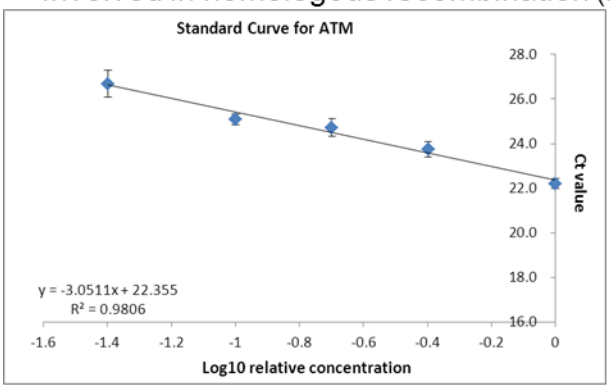


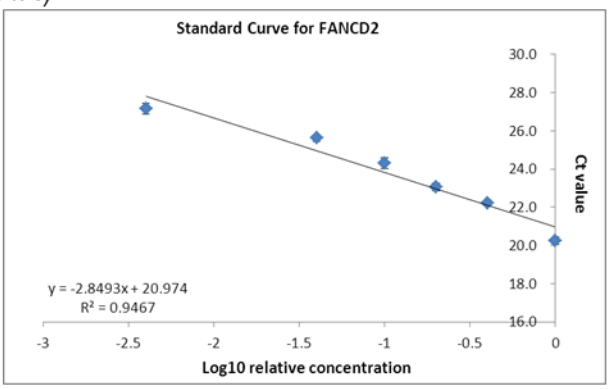


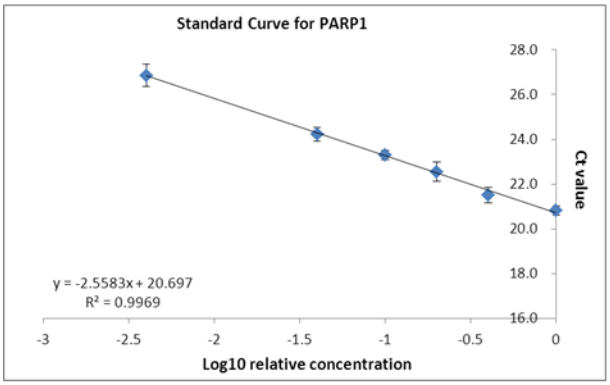


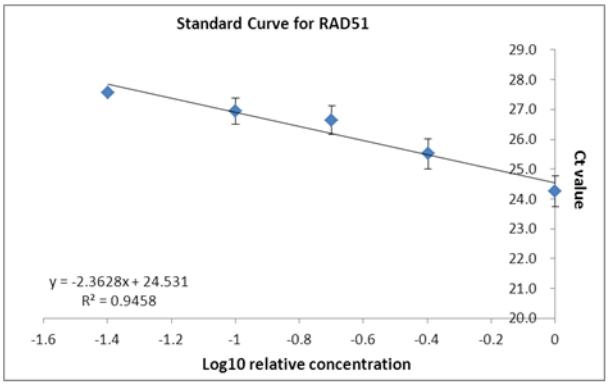
C) Genes involved in DNA damage responses and genes encoding MERIT40 interacting proteints within the BRCA1 complex

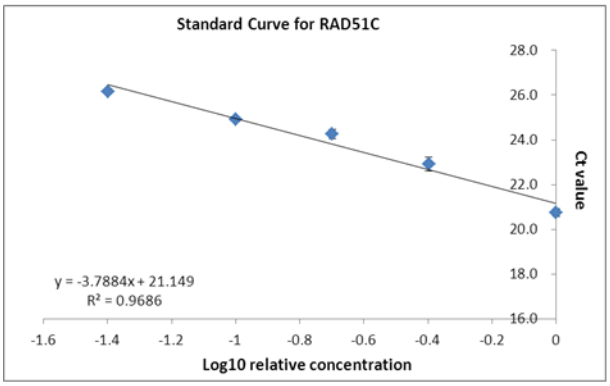
- Involved in homologous recombination (HR)

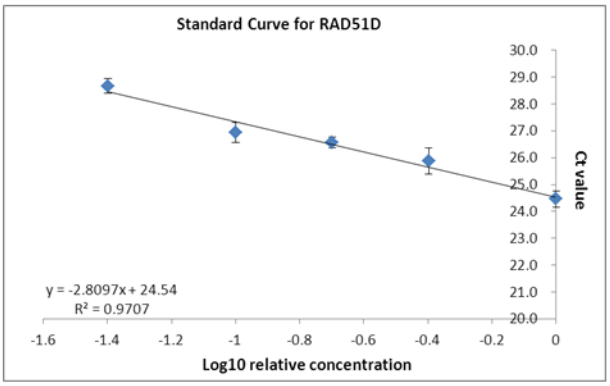


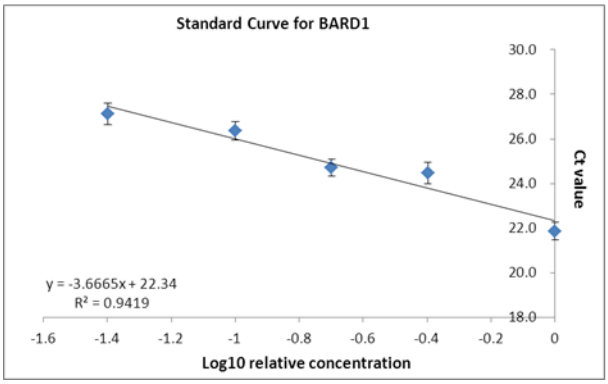


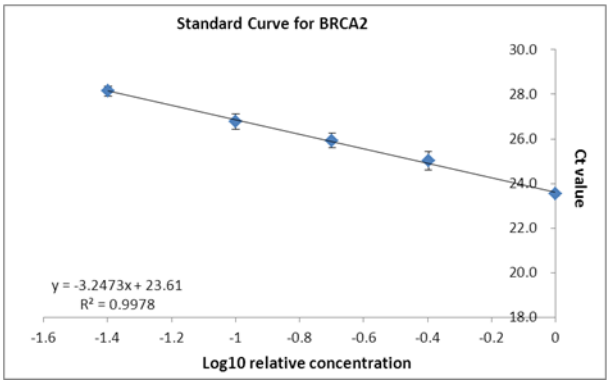




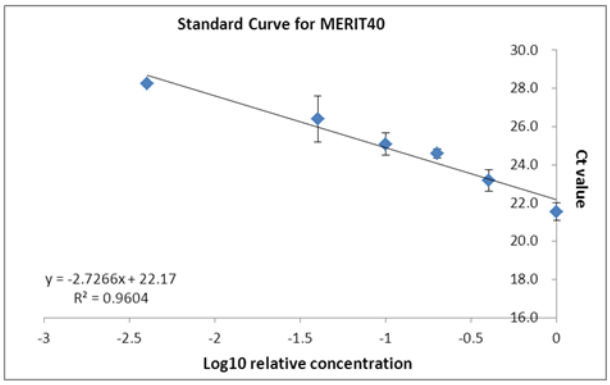


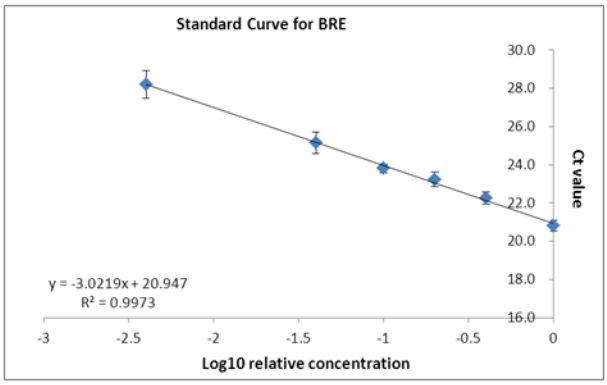


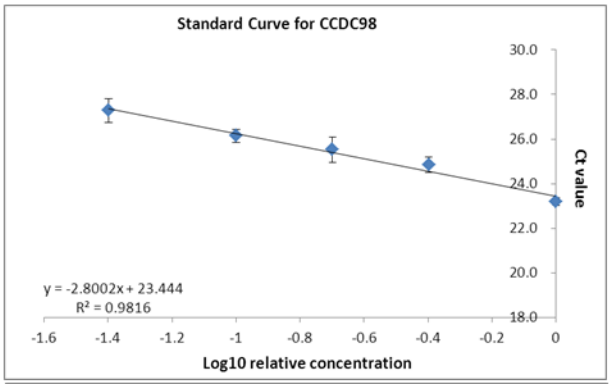


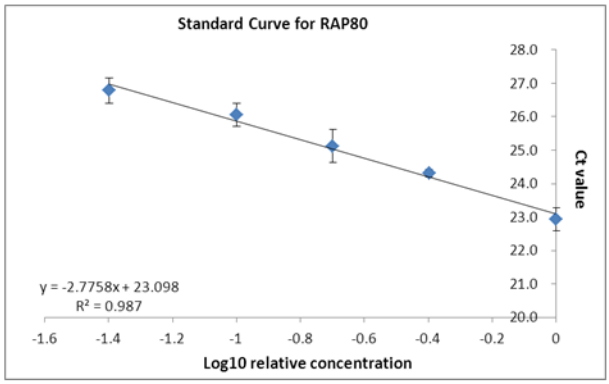


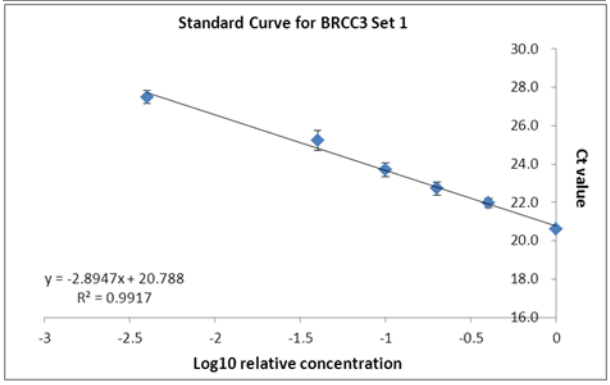
- Involved in HR as part of the BRCA1 interacting complex along with MERIT40

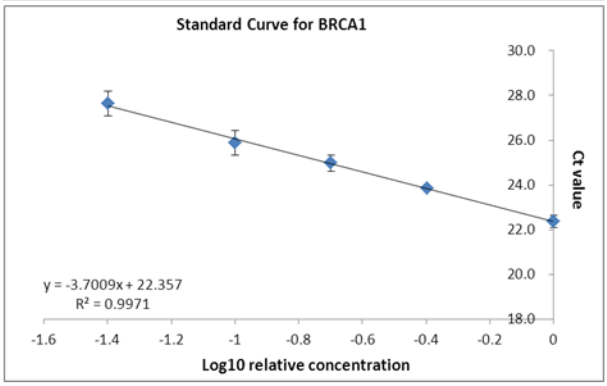




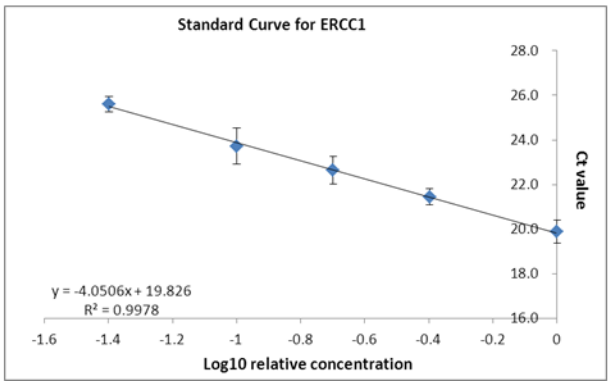


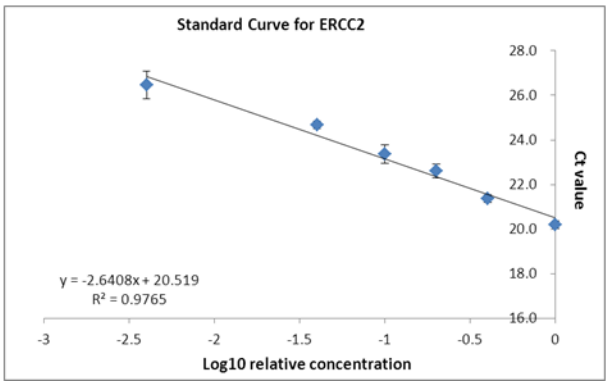


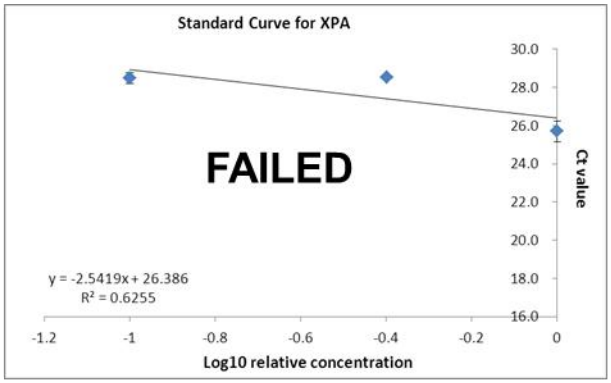


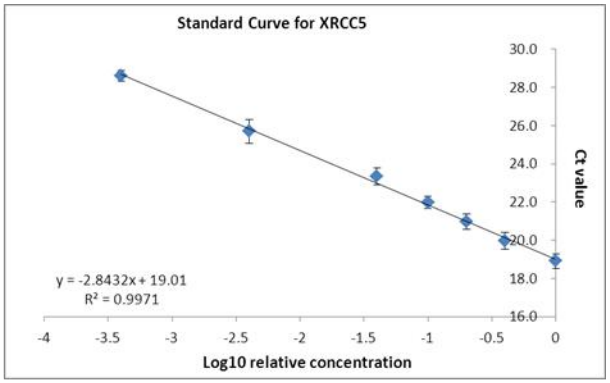


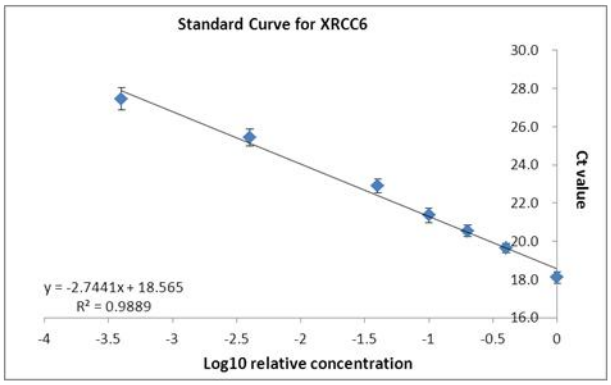
- Involved in nucleotide excision repair (NER)

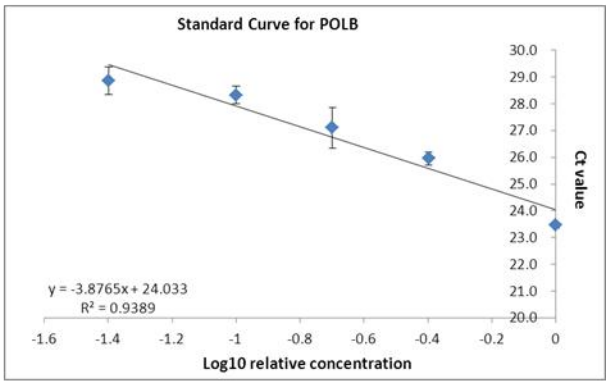




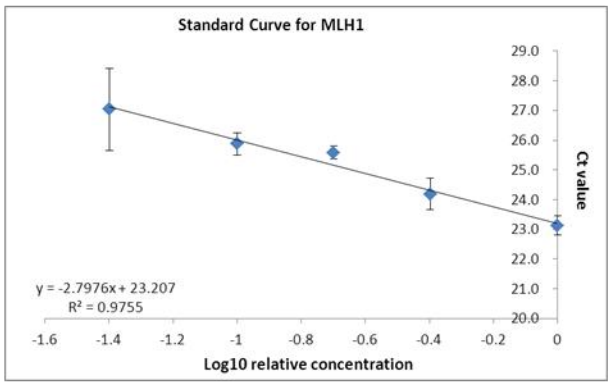


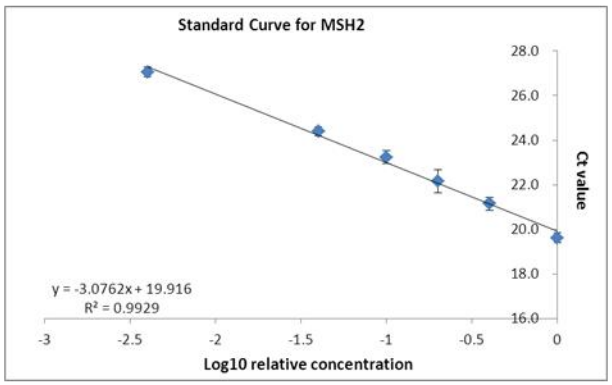


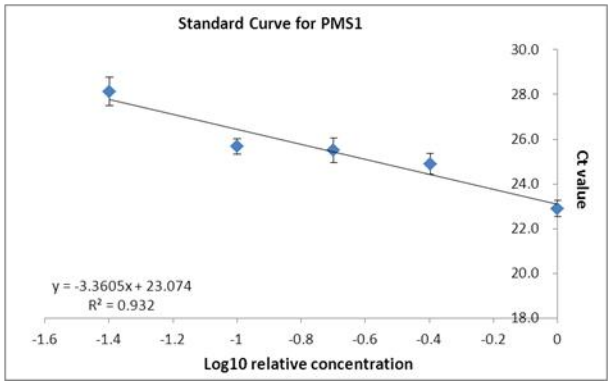




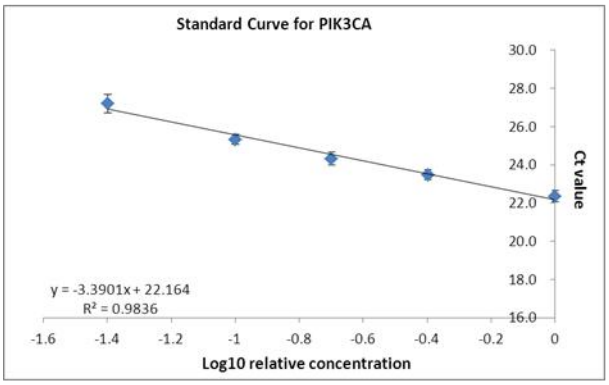
- Involved in mismatch repair (MR)

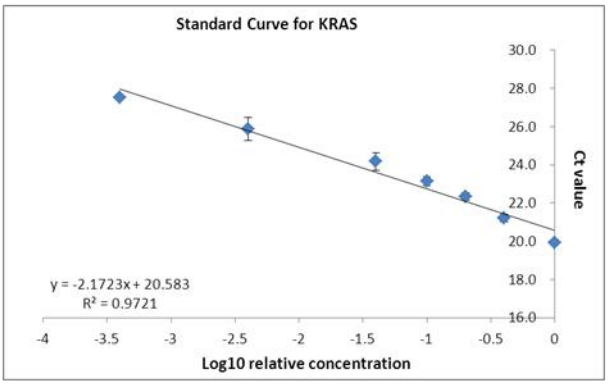


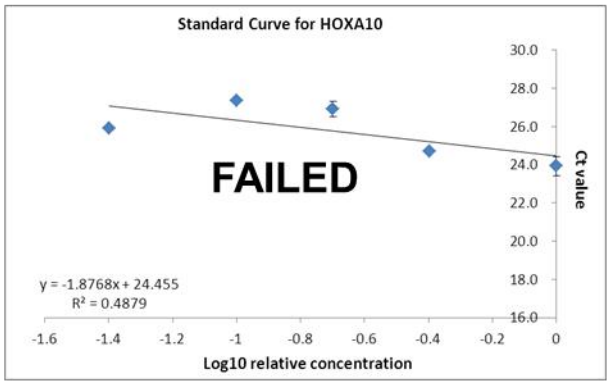


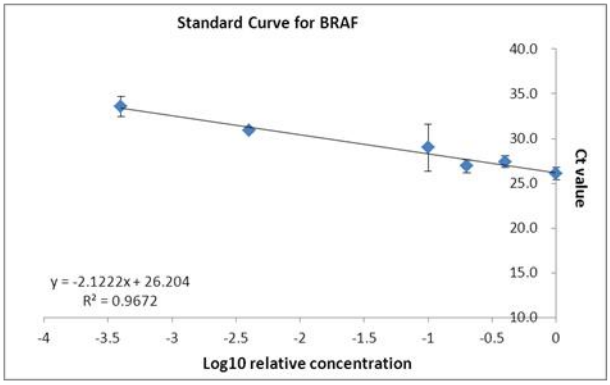


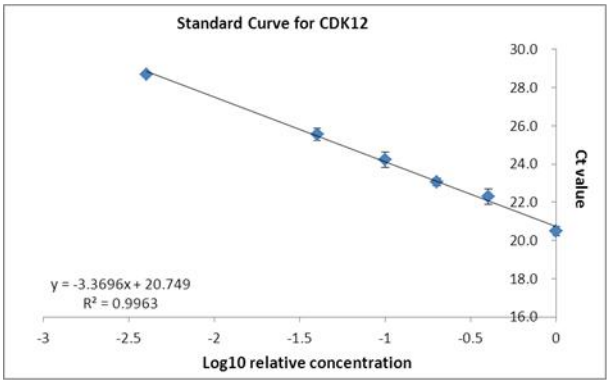
C) Genes involved in cell proliferation, differentiation and apoptotic pathways

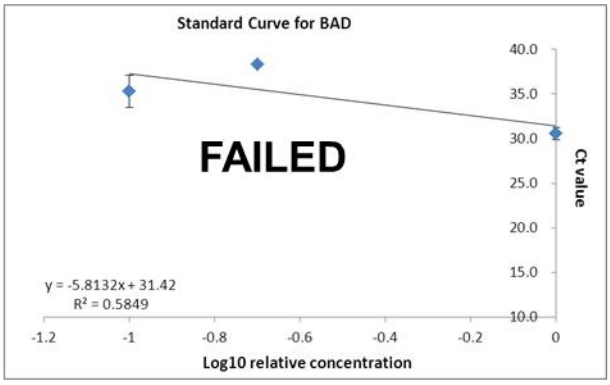


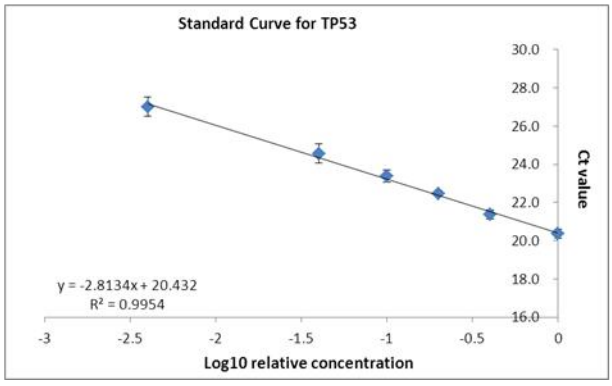




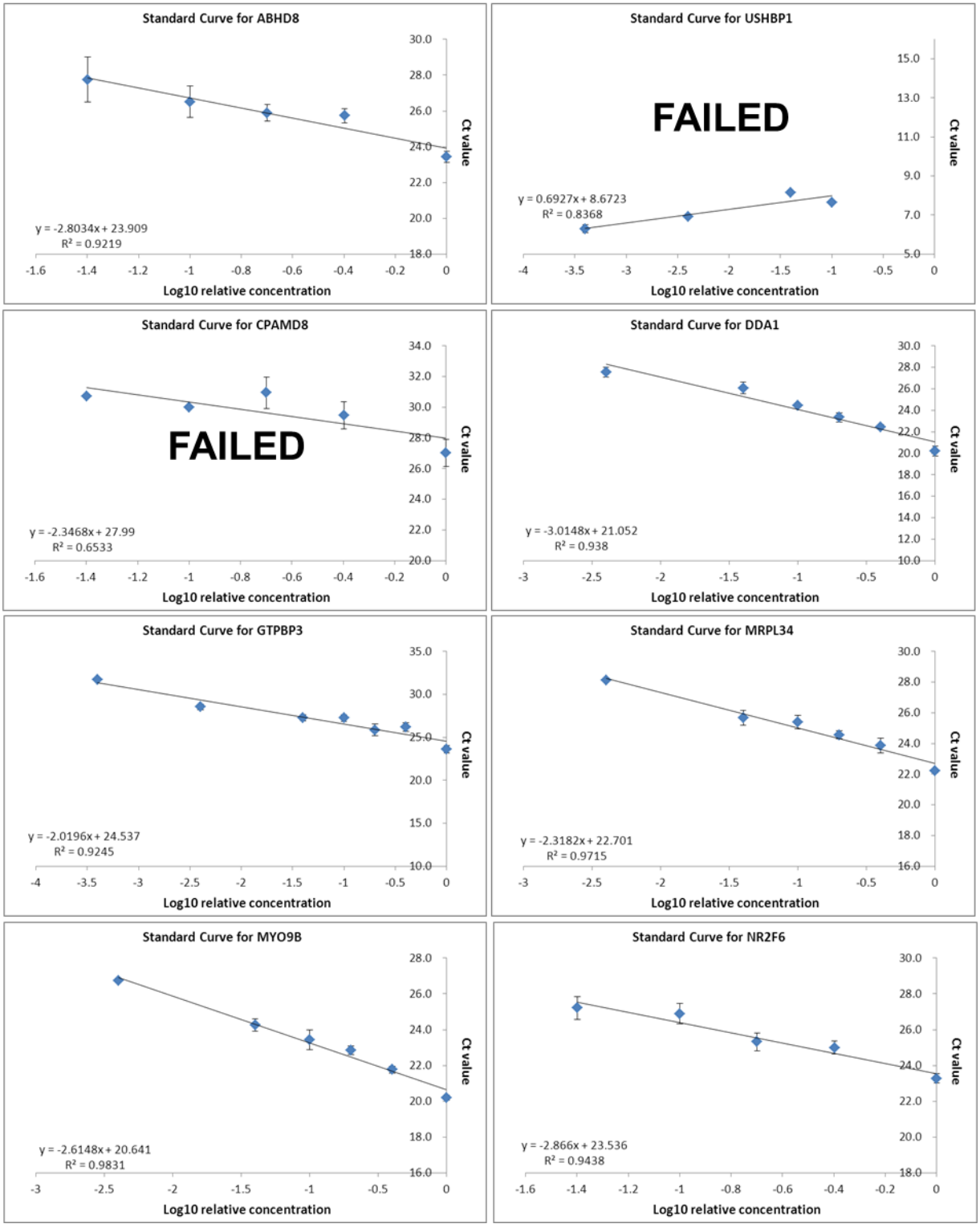


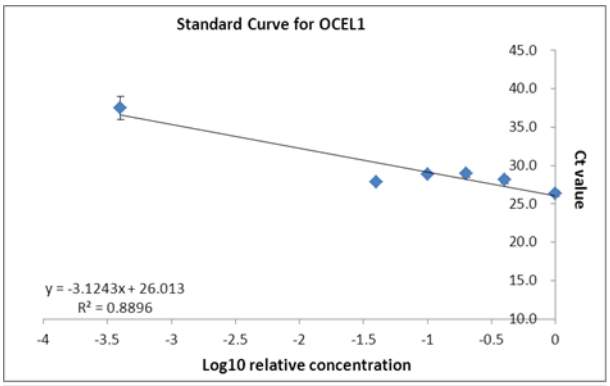


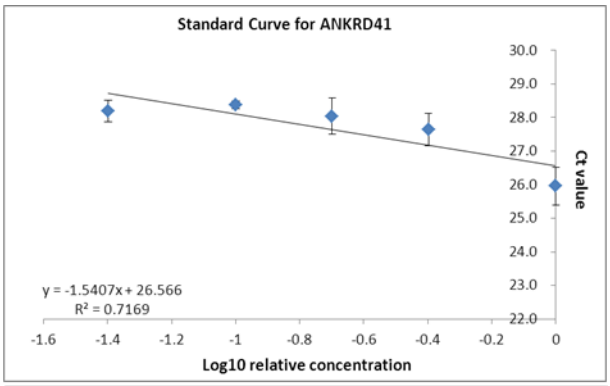


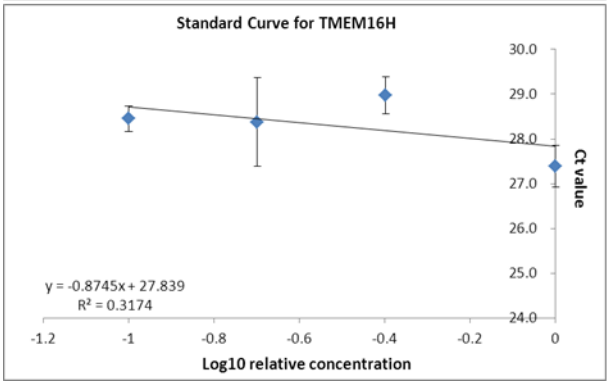


C) Genes in chromosome 19 in close proximity to MERIT40




Standard Curve for OCEL1
y = -3.1243x + 26.013
R² = 0.8896
Ct value
Log10 relative concentration


Standard Curve for ANKRD41
y = -1.5407x + 26.566
R² = 0.7169
Ct value
Log10 relative concentration


Standard Curve for TMEM16H
y = -0.8745x + 27.839
R² = 0.3174
Ct value
Log10 relative concentration

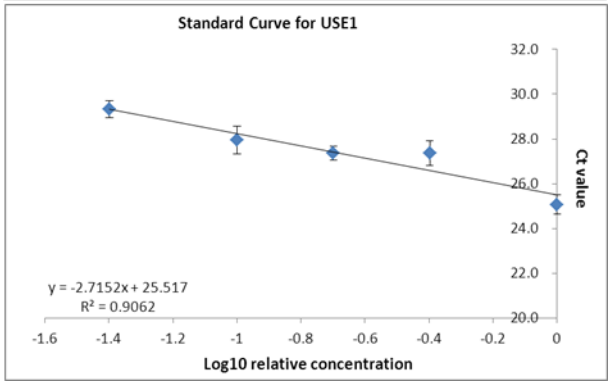

Standard Curve for USE1
y = -2.7152x + 25.517
R² = 0.9062
Ct value
Log10 relative concentration

Figure 4: Close inspection of cisplatin dose response curve for A2780CP control and *MERIT40* knockdown cell lines. The dose response curve for increasing doses of cisplatin up to 150µM did not show any differences in the cell lines' cisplatin sensitivity. Close inspection of the curve clearly shows the difference of the A2780CP _M1 cells response up to 7.5um of cisplatin dosing. The error bars are showing the SEM (standard error of mean) which is the standard error of the mean representing how accurate the estimate of the mean cell viability is for each cisplatin concentration.

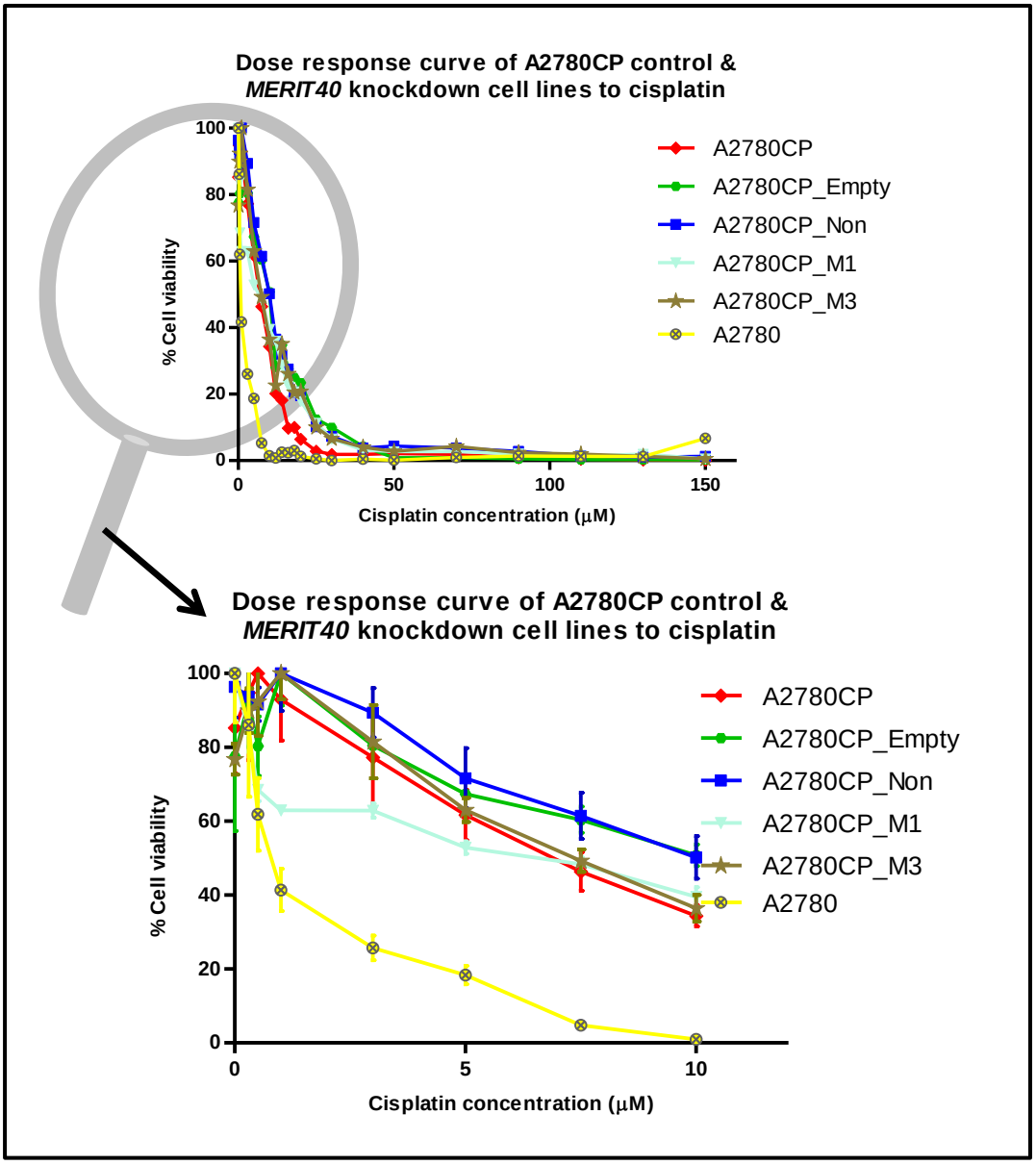


Figure 5: Differential expression of *PARP1* in EOC and normal cell lines

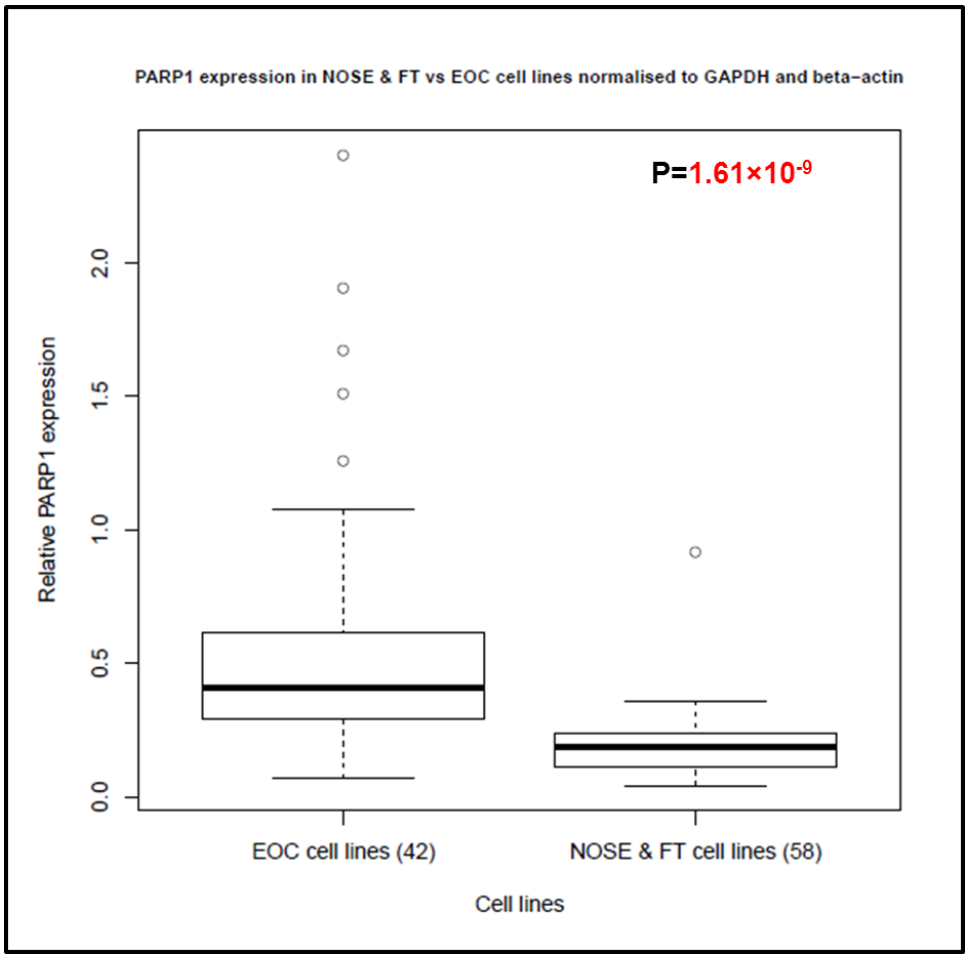


Figure 6: *MSH2* expression in normal and EOC cell lines A) Differential expression of *MSH2* between EOC and normal cell lines, B) Differential expression of *MSH2* between the 4 EOC cell lines selected to knock down *MERIT40*

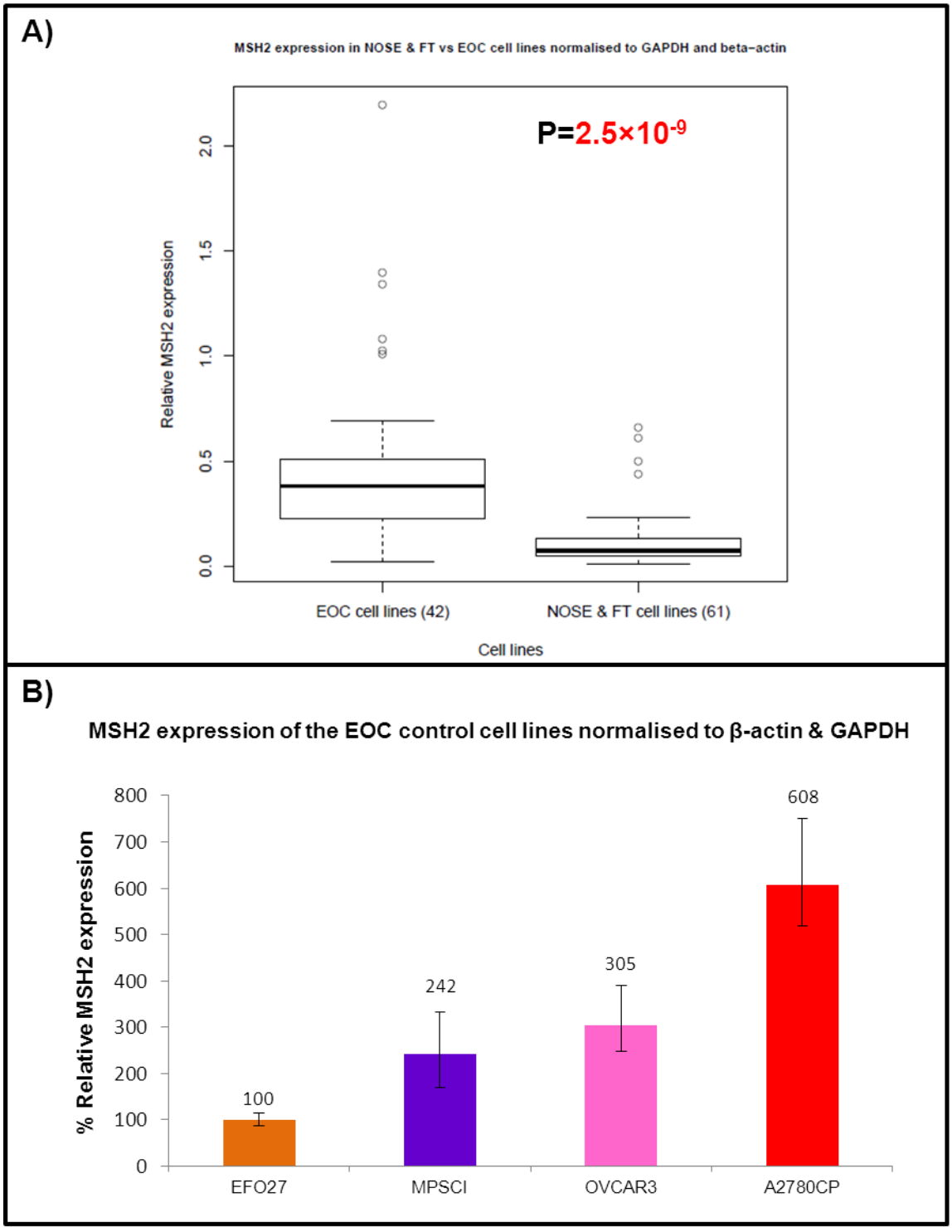


Table 1: Summary of relative expression (%) of selected genes in control and *MERIT40* knockdown EFO27 cell lines. The confidence intervals were calculated from +/- error of relative expression.

Gene	Fluidigm Chip	Normalised against	% Relative expresion [Confidence intervals]			
			EFO27			
			Empty	Non-silencing	M1	M3
CCNE1	Chip 1	β-actin & GAPDH	100 [116-86]	190 [222-165]	185 [234-152]	203 [251-165]
CDKN1A	Chip 5	β-actin & GAPDH	100 [125-80]	121 [148-99]	65 [81-53]	112 [147-85]
CDKN2A	Chip 5	β-actin & GAPDH	100 [110-91]	140 [166-118]	106 [164-71]	151 [186-123]
PTEN	Chip 5	β-actin & GAPDH	100 [117-86]	93 [129-67]	104 [135-82]	117 [136-101]
RB1	Chip 7	β-actin & GAPDH	100 [129-77]	101 [123-83]	85 [101-72]	80 [89-72]
BUBR1	Chip 4	β-actin & GAPDH	100 [113-89]	103 [115-93]	97 [117-82]	95 [115-79]
MAD1	Chip 5	β-actin & GAPDH	100 [122-82]	99 [122-81]	57 [73-45]	69 [117-41]
MAD2L1	Chip 5	β-actin & GAPDH	100 [147-68]	160 [202-127]	92 [131-67]	114 [167-78]
MAD2L2	Chip 5	β-actin & GAPDH	100 [112-89]	119 [137-104]	116 [147-93]	120 [153-94]
CHK1	Chip 5	β-actin & GAPDH	100 [113-89]	122 [136-110]	112 [138-92]	112 [126-99]
CHK2	Chip 5	β-actin & GAPDH	100 [113-88]	153 [193-122]	76 [153-40]	154 [217-109]
BUB1	Chip 4	β-actin & GAPDH	100 [134-75]	122 [147-102]	104 [144-77]	122 [148-101]
BUB3	Chip 4	β-actin & GAPDH	100 [112-89]	96 [112-83]	97 [122-79]	110 [135-90]
ATM	Chip 4	β-actin & GAPDH	100 [124-81]	80 [95-68]	99 [123-81]	88 [109-72]
FANCD2	Chip 5	β-actin & GAPDH	100 [118-85]	150 [176-128]	154 [196-124]	123 [162-94]
PARP1	Chip 5	β-actin & GAPDH	100 [131-76]	157 [199-124]	123 [150-102]	150 [181-125]
RAD51	Chip 5	β-actin & GAPDH	100 [117-86]	174 [209-145]	135 [188-99]	110 [139-87]
RAD51C	Chip 5	β-actin & GAPDH	100 [153-65]	194 [281-134]	138 [212-93]	127 [189-85]
RAD51D	Chip 5	β-actin & GAPDH	100 [138-73]	157 [194-127]	166 [221-128]	150 [174-129]
BARD1	Chip 5	β-actin & GAPDH	100 [115-87]	83 [115-60]	103 [144-75]	190 [245-147]
BRCA1	Chip 4	β-actin & GAPDH	100 [121-83]	106 [120-94]	92 [115-75]	101 [152-68]
BRCA2	Chip 4	β-actin & GAPDH	100 [138-72]	89 [100-80]	104 [129-85]	100 [141-72]
MERIT40	Chip 2	β-actin	100 [113-89]	165 [185-147]	11 [16-7]	49 [60-40]
BRE	Chip 4	β-actin & GAPDH	100 [116-86]	76 [98-60]	89 [106-76]	98 [123-79]
CCDC98	Chip 4	β-actin & GAPDH	100 [126-79]	97 [123-77]	60 [73-50]	96 [116-80]
RAP80	Chip 5	β-actin & GAPDH	100 [127-79]	168 [188-151]	122 [189-82]	136 [160-115]
BRCC3	Chip 1	β-actin & GAPDH	100 [109-92]	156 [189-130]	105 [172-69]	129 [166-100]
ERCC1	Chip 5	β-actin & GAPDH	100 [140-71]	109 [160-75]	94 [143-64]	132 [162-107]
ERCC2	Chip 5	β-actin & GAPDH	100 [111-90]	143 [169-121]	139 [176-112]	129 [155-108]
XRCC5	Chip 5	β-actin & GAPDH	100 [132-76]	133 [176-101]	94 [126-72]	125 [156-100]
XRCC6	Chip 5	β-actin & GAPDH	100 [119-84]	132 [164-106]	126 [183-89]	120 [162-89]
POLB	Chip 5	β-actin & GAPDH	100 [138-72]	181 [213-154]	234 [284-196]	164 [226-119]
MLH1	Chip 5	β-actin & GAPDH	100 [112-89]	181 [212-155]	134 [171-107]	155 [187-128]
MSH2	Chip 5	β-actin & GAPDH	100 [114-87]	138 [180-106]	125 [177-91]	122 [173-86]
PMS1	Chip 5	β-actin & GAPDH	100 [178-56]	282 [525-152]	289 [403-213]	118 [136-103]
PIK3CA	Chip 5	β-actin & GAPDH	100 [115-87]	143 [186-110]	132 [163-109]	96 [112-82]
KRAS	Chip 5	β-actin & GAPDH	100 [132-76]	153 [187-126]	133 [175-103]	122 [162-92]
BRAF	Chip 4	β-actin & GAPDH	100 [120-83]	60 [90-41]	108 [143-83]	112 [157-81]
CDK12	Chip 6	β-actin & GAPDH	100 [116-86]	76 [89-65]	54 [66-45]	72 [94-54]
TP53	Chip 5	β-actin & GAPDH	100 [112-89]	104 [117-93]	100 [128-80]	96 [119-77]
ABHD8	Chip 2	β-actin	100 [143-70]	131 [146-117]	92 [128-66]	106 [131-86]
DDA1	Chip 2	β-actin	100 [117-85]	111 [130-95]	112 [149-84]	127 [165-98]
GTPBP3	Chip 2	β-actin	100 [128-78]	82 [113-59]	126 [139-114]	181 [266-123]
MRPL34	Chip 3	β-actin & GAPDH	100 [112-89]	95 [124-74]	111 [134-95]	140 [151-130]
MYO9B	Chip 3	β-actin & GAPDH	100 [103-97]	118 [135-104]	97 [102-93]	95 [101-89]
NR2F6	Chip 3	β-actin & GAPDH	100 [112-90]	65 [99-44]	85 [111-67]	139 [149-129]
USE1	Chip 4	β-actin & GAPDH	100 [150-67]	123 [155-99]	163 [231-119]	183 [222-152]
β-actin SET 1	Chip 1	GAPDH	100 [115-87]	114 [140-93]	151 [183-124]	103 [129-83]
GAPDH SET 1	Chip 1	β-actin	100 [115-87]	87 [107-71]	66 [80-54]	97 [121-78]
GAPDH SET 2	Chip 2	N/A (β-actin failed)	N/A	N/A	N/A	N/A
β-actin SET 3	Chip 3	GAPDH	100 [110-91]	120 [135-107]	143 [150-136]	95 [106-85]
GAPDH SET 3	Chip 3	β-actin	100 [110-91]	83 [93-74]	70 [73-67]	105 [117-94]
β-actin SET 4	Chip 4	GAPDH	100 [118-85]	113 [128-100]	123 [150-101]	109 [141-84]
GAPDH SET 4	Chip 4	β-actin	100 [118-85]	89 [101-79]	81 [99-66]	92 [119-71]
β-actin SET 5	Chip 5	GAPDH	100 [122-82]	102 [134-77]	121 [160-91]	100 [130-77]
GAPDH SET 5	Chip 5	β-actin	100 [122-82]	98 [129-75]	83 [110-63]	100 [130-77]
β-actin SET 6	Chip 6	GAPDH	100 [122-82]	99 [130-76]	144 [158-131]	76 [95-61]
GAPDH SET 6	Chip 6	β-actin	100 [122-82]	101 [132-77]	70 [77-64]	132 [165-105]
GAPDH SET 7	Chip 7	N/A (β-actin failed)	N/A	N/A	N/A	N/A

Table 2: Summary of relative expression (%) of selected genes in control and *MERIT40* knockdown MPSC1 cell lines. The confidence intervals were calculated from +/- error of relative expression.

Gene	Fluidigm Chip	Normalised against	% Relative expresion [Confidence intervals]			
			MPSC1			
			Empty	Non-silencing	M1	M3
CCNE1	Chip 1	β-actin & GAPDH	100 [123-81]	88 [131-59]	73 [88-62]	94 [114-78]
CDKN1A	Chip 5	β-actin & GAPDH	100 [135-74]	97 [127-74]	86 [95-77]	96 [111-83]
CDKN2A	Chip 5	β-actin & GAPDH	FAILED	FAILED	FAILED	FAILED
PTEN	Chip 5	β-actin & GAPDH	100 [129-78]	86 [117-63]	81 [98-67]	92 [119-71]
RB1	Chip 7	β-actin & GAPDH	100 [115-87]	99 [108-91]	101 [130-79]	99 [129-76]
BUBR1	Chip 4	β-actin & GAPDH	100 [110-91]	109 [134-89]	64 [86-48]	127 [145-111]
MAD1	Chip 5	β-actin & GAPDH	100 [141-71]	91 [109-76]	95 [114-79]	124 [172-90]
MAD2L1	Chip 5	β-actin & GAPDH	100 [140-71]	103 [140-76]	83 [101-68]	123 [169-90]
MAD2L2	Chip 5	β-actin & GAPDH	100 [140-72]	105 [134-82]	93 [116-75]	134 [150-119]
CHK1	Chip 5	β-actin & GAPDH	100 [137-73]	122 [139-107]	96 [111-83]	152 [171-135]
CHK2	Chip 5	β-actin & GAPDH	100 [129-78]	177 [253-124]	111 [163-76]	146 [204-105]
BUB1	Chip 4	β-actin & GAPDH	100 [111-90]	126 [172-92]	58 [73-46]	124 [146-105]
BUB3	Chip 4	β-actin & GAPDH	100 [123-82]	130 [161-105]	72 [84-61]	117 [133-103]
ATM	Chip 4	β-actin & GAPDH	100 [116-86]	152 [217-106]	93 [131-66]	109 [158-75]
FANCD2	Chip 5	β-actin & GAPDH	100 [159-63]	93 [129-67]	103 [114-93]	129 [152-110]
PARP1	Chip 5	β-actin & GAPDH	100 [141-71]	101 [114-89]	96 [129-71]	142 [178-113]
RAD51	Chip 5	β-actin & GAPDH	100 [136-73]	113 [129-99]	96 [127-72]	101 [115-89]
RAD51C	Chip 5	β-actin & GAPDH	100 [147-68]	69 [78-61]	86 [165-45]	167 [237-117]
RAD51D	Chip 5	β-actin & GAPDH	100 [150-67]	78 [101-60]	72 [85-61]	88 [108-71]
BARD1	Chip 5	β-actin & GAPDH	100 [132-76]	76 [87-66]	86 [102-73]	89 [107-74]
BRCA1	Chip 4	β-actin & GAPDH	100 [119-84]	117 [146-94]	62 [71-54]	104 [116-93]
BRCA2	Chip 4	β-actin & GAPDH	100 [113-89]	135 [176-104]	89 [131-60]	100 [128-78]
MERIT40	Chip 2	β-actin	100 [110-91]	107 [129-88]	12 [16-9]	51 [62-42]
BRE	Chip 4	β-actin & GAPDH	100 [132-76]	94 [132-67]	46 [66-32]	65 [83-51]
CCDC98	Chip 4	β-actin & GAPDH	100 [140-72]	107 [166-69]	70 [90-55]	93 [107-81]
RAP80	Chip 5	β-actin & GAPDH	100 [129-78]	130 [172-99]	96 [118-78]	168 [187-151]
BRCC3	Chip 1	β-actin & GAPDH	100 [126-79]	143 [191-107]	100 [135-74]	99 [143-68]
ERCC1	Chip 5	β-actin & GAPDH	100 [130-77]	81 [116-56]	71 [108-47]	66 [100-43]
ERCC2	Chip 5	β-actin & GAPDH	100 [131-77]	98 [113-85]	101 [112-91]	125 [141-111]
XRCC5	Chip 5	β-actin & GAPDH	100 [147-68]	107 [133-86]	96 [132-70]	141 [176-113]
XRCC6	Chip 5	β-actin & GAPDH	100 [133-75]	110 [142-85]	107 [136-84]	157 [196-126]
POLB	Chip 5	β-actin & GAPDH	100 [135-74]	95 [109-82]	83 [102-67]	124 [164-94]
MLH1	Chip 5	β-actin & GAPDH	100 [136-74]	101 [120-85]	103 [114-93]	103 [132-80]
MSH2	Chip 5	β-actin & GAPDH	100 [142-70]	103 [143-74]	84 [95-74]	149 [196-114]
PMS1	Chip 5	β-actin & GAPDH	100 [155-65]	90 [116-70]	138 [161-118]	144 [163-127]
PIK3CA	Chip 5	β-actin & GAPDH	100 [134-75]	99 [112-88]	73 [103-52]	91 [119-69]
KRAS	Chip 5	β-actin & GAPDH	100 [135-74]	98 [112-86]	79 [92-68]	93 [104-83]
BRAF	Chip 4	β-actin & GAPDH	100 [109-91]	116 [143-94]	51 [65-40]	40 [52-31]
CDK12	Chip 6	β-actin & GAPDH	100 [148-68]	116 [152-89]	106 [139-81]	106 [146-77]
TP53	Chip 5	β-actin & GAPDH	100 [134-75]	109 [125-95]	108 [123-95]	106 [148-76]
ABHD8	Chip 2	β-actin	100 [122-82]	124 [154-100]	83 [105-65]	55 [60-50]
DDA1	Chip 2	β-actin	100 [122-82]	125 [155-101]	91 [110-75]	60 [81-44]
GTPBP3	Chip 2	β-actin	100 [109-92]	128 [190-86]	95 [115-79]	79 [86-73]
MRPL34	Chip 3	β-actin & GAPDH	100 [112-90]	311 [411-235]	115 [139-95]	71 [93-54]
MYO9B	Chip 3	β-actin & GAPDH	100 [114-88]	129 [167-99]	115 [148-89]	78 [96-63]
NR2F6	Chip 3	β-actin & GAPDH	100 [118-85]	190 [250-144]	108 [132-88]	78 [105-58]
USE1	Chip 4	β-actin & GAPDH	100 [106-94]	88 [131-59]	45 [65-31]	88 [118-66]
β-actin SET 1	Chip 1	GAPDH	100 [129-77]	87 [128-59]	111 [128-97]	93 [118-73]
GAPDH SET 1	Chip 1	β-actin	100 [129-77]	115 [169-78]	90 [103-78]	107 [135-85]
GAPDH SET 2	Chip 2	N/A (β-actin failed)	N/A	N/A	N/A	N/A
β-actin SET 3	Chip 3	GAPDH	100 [123-81]	71 [103-49]	102 [135-77]	106 [143-79]
GAPDH SET 3	Chip 3	β-actin	100 [123-81]	141 [203-98]	98 [130-74]	94 [127-70]
β-actin SET 4	Chip 4	GAPDH	100 [137-73]	118 [170-82]	199 [238-166]	149 [170-131]
GAPDH SET 4	Chip 4	β-actin	100 [137-73]	84 [121-58]	50 [60-42]	67 [76-59]
β-actin SET 5	Chip 5	GAPDH	100 [144-70]	89 [110-72]	74 [91-60]	83 [104-66]
GAPDH SET 5	Chip 5	β-actin	100 [144-70]	113 [139-92]	136 [167-111]	121 [151-97]
β-actin SET 6	Chip 6	GAPDH	100 [147-68]	101 [143-71]	113 [142-90]	118 [151-92]
GAPDH SET 6	Chip 6	β-actin	100 [147-68]	99 [140-70]	89 [112-71]	85 [109-67]
GAPDH SET 7	Chip 7	N/A (β-actin failed)	N/A	N/A	N/A	N/A

Table 3: Summary of relative expression (%) of selected genes in control and *MERIT40* knockdown OVCAR3 cell lines. The confidence intervals were calculated from +/- error of relative expression.

Gene	Fluidigm Chip	Normalised against	% Relative expresion [Confidence intervals]			
			OVCAR3			
			Empty	Non-silencing	M1	M3
CCNE1	Chip 1	β-actin & GAPDH	100 [111-90]	89 [112-70]	114 [125-104]	65 [83-51]
CDKN1A	Chip 5	β-actin & GAPDH	100 [116-86]	87 [127-60]	65 [81-52]	50 [64-39]
CDKN2A	Chip 5	β-actin & GAPDH	100 [119-84]	74 [107-51]	47 [67-33]	58 [79-43]
PTEN	Chip 5	β-actin & GAPDH	100 [141-71]	93 [135-64]	63 [73-55]	100 [119-84]
RB1	Chip 7	β-actin & GAPDH	100 [116-86]	89 [101-79]	78 [93-66]	79 [97-65]
BUBR1	Chip 4	β-actin & GAPDH	100 [119-84]	93 [104-83]	114 [134-97]	155 [193-125]
MAD1	Chip 5	β-actin & GAPDH	100 [146-68]	133 [198-89]	74 [84-65]	103 [134-79]
MAD2L1	Chip 5	β-actin & GAPDH	100 [127-78]	86 [135-55]	89 [116-68]	101 [131-78]
MAD2L2	Chip 5	β-actin & GAPDH	100 [125-80]	80 [121-53]	75 [101-56]	84 [105-67]
CHK1	Chip 5	β-actin & GAPDH	100 [114-88]	100 [146-68]	82 [96-70]	111 [136-91]
CHK2	Chip 5	β-actin & GAPDH	FAILED	FAILED	FAILED	FAILED
BUB1	Chip 4	β-actin & GAPDH	100 [118-85]	75 [97-58]	68 [79-58]	105 [125-88]
BUB3	Chip 4	β-actin & GAPDH	100 [112-89]	92 [100-84]	89 [100-79]	136 [161-115]
ATM	Chip 4	β-actin & GAPDH	100 [113-89]	126 [145-110]	102 [118-88]	152 [193-120]
FANCD2	Chip 5	β-actin & GAPDH	100 [112-89]	100 [147-68]	50 [58-43]	79 [95-66]
PARP1	Chip 5	β-actin & GAPDH	100 [125-80]	115 [169-78]	61 [71-53]	82 [110-61]
RAD51	Chip 5	β-actin & GAPDH	100 [118-85]	98 [151-64]	99 [114-86]	84 [119-59]
RAD51C	Chip 5	β-actin & GAPDH	100 [131-76]	78 [121-50]	92 [125-68]	120 [179-81]
RAD51D	Chip 5	β-actin & GAPDH	100 [129-78]	127 [189-85]	45 [51-40]	105 [131-84]
BARD1	Chip 5	β-actin & GAPDH	100 [111-90]	67 [100-45]	103 [120-88]	77 [98-61]
BRCA1	Chip 4	β-actin & GAPDH	100 [115-87]	96 [126-73]	118 [150-93]	163 [203-131]
BRCA2	Chip 4	β-actin & GAPDH	100 [116-86]	86 [109-68]	99 [112-88]	109 [133-89]
MERIT40	Chip 2	β-actin	100 [119-84]	86 [103-72]	11 [16-8]	61 [80-46]
BRE	Chip 4	β-actin & GAPDH	100 [122-82]	87 [99-77]	80 [94-68]	153 [186-126]
CCDC98	Chip 4	β-actin & GAPDH	100 [125-80]	92 [106-80]	49 [62-39]	126 [146-109]
RAP80	Chip 5	β-actin & GAPDH	100 [133-75]	82 [124-54]	63 [73-54]	77 [101-59]
BRCC3	Chip 1	β-actin & GAPDH	100 [130-77]	84 [88-80]	95 [114-79]	94 [113-78]
ERCC1	Chip 5	β-actin & GAPDH	100 [148-68]	55 [86-35]	41 [60-28]	55 [81-37]
ERCC2	Chip 5	β-actin & GAPDH	100 [112-90]	117 [169-81]	75 [89-64]	94 [120-74]
XRCC5	Chip 5	β-actin & GAPDH	100 [123-81]	80 [124-52]	62 [75-51]	96 [122-76]
XRCC6	Chip 5	β-actin & GAPDH	100 [130-77]	125 [190-82]	70 [84-58]	104 [132-82]
POLB	Chip 5	β-actin & GAPDH	100 [122-82]	125 [192-81]	77 [92-64]	97 [134-70]
MLH1	Chip 5	β-actin & GAPDH	FAILED	FAILED	FAILED	FAILED
MSH2	Chip 5	β-actin & GAPDH	100 [123-81]	144 [213-97]	91 [109-76]	113 [140-91]
PMS1	Chip 5	β-actin & GAPDH	100 [130-77]	34 [54-21]	59 [76-46]	76 [116-50]
PIK3CA	Chip 5	β-actin & GAPDH	100 [129-77]	91 [133-62]	56 [71-44]	78 [101-60]
KRAS	Chip 5	β-actin & GAPDH	100 [132-75]	84 [121-58]	62 [73-52]	75 [89-64]
BRAF	Chip 4	β-actin & GAPDH	100 [152-66]	109 [136-87]	145 [171-123]	125 [179-87]
CDK12	Chip 6	β-actin & GAPDH	100 [120-83]	107 [151-76]	83 [107-64]	109 [139-86]
TP53	Chip 5	β-actin & GAPDH	100 [115-87]	94 [136-65]	62 [76-51]	73 [94-57]
ABHD8	Chip 2	β-actin	100 [120-83]	89 [110-72]	38 [45-32]	62 [78-49]
DDA1	Chip 2	β-actin	100 [129-78]	73 [83-64]	122 [137-109]	155 [201-119]
GTPBP3	Chip 2	β-actin	100 [113-89]	40 [56-29]	209 [233-187]	305 [379-245]
MRPL34	Chip 3	β-actin & GAPDH	100 [126-80]	103 [110-97]	196 [230-167]	426 [472-385]
MYO9B	Chip 3	β-actin & GAPDH	100 [123-81]	105 [113-98]	82 [98-69]	102 [111-93]
NR2F6	Chip 3	β-actin & GAPDH	100 [120-83]	79 [82-76]	199 [245-162]	348 [392-309]
USE1	Chip 4	β-actin & GAPDH	100 [139-72]	65 [75-56]	105 [129-85]	90 [102-80]
β-actin SET 1	Chip 1	GAPDH	100 [113-89]	97 [104-90]	65 [78-54]	103 [130-82]
GAPDH SET 1	Chip 1	β-actin	100 [113-89]	103 [110-96]	154 [184-129]	97 [122-77]
GAPDH SET 2	Chip 2	N/A (β-actin failed)	N/A	N/A	N/A	N/A
β-actin SET 3	Chip 3	GAPDH	100 [127-79]	88 [94-83]	79 [99-63]	87 [103-73]
GAPDH SET 3	Chip 3	β-actin	100 [127-79]	114 [122-107]	127 [160-101]	114 [135-96]
β-actin SET 4	Chip 4	GAPDH	100 [120-83]	79 [89-70]	64 [70-58]	60 [67-54]
GAPDH SET 4	Chip 4	β-actin	100 [120-83]	126 [141-112]	157 [173-142]	168 [187-151]
β-actin SET 5	Chip 5	GAPDH	100 [120-83]	74 [125-44]	62 [81-47]	100 [130-77]
GAPDH SET 5	Chip 5	β-actin	100 [120-83]	134 [226-79]	160 [209-123]	100 [130-77]
β-actin SET 6	Chip 6	GAPDH	100 [125-80]	93 [129-67]	73 [94-57]	89 [102-77]
GAPDH SET 6	Chip 6	β-actin	100 [125-80]	108 [150-78]	136 [175-106]	113 [130-98]
GAPDH SET 7	Chip 7	N/A (β-actin failed)	N/A	N/A	N/A	N/A

Table 4: Summary of relative expression (%) of selected genes in control and *MERIT40* knockdown A2780CP cell lines. The confidence intervals were calculated from +/- error of relative expression.

Gene	Fluidigm Chip	Normalised against	% Relative expresion [Confidence intervals]			
			A2780CP			
			Empty	Non-silencing	M1	M3
CCNE1	Chip 1	β-actin & GAPDH	100 [123-81]	97 [127-74]	101 [118-87]	130 [167-102]
CDKN1A	Chip 5	β-actin & GAPDH	100 [158-63]	19 [26-14]	123 [153-99]	41 [56-30]
CDKN2A	Chip 5	β-actin & GAPDH	100 [154-65]	76 [92-63]	97 [105-90]	58 [86-39]
PTEN	Chip 5	β-actin & GAPDH	100 [143-70]	66 [70-63]	83 [90-76]	71 [85-59]
RB1	Chip 7	β-actin & GAPDH	100 [122-82]	80 [102-63]	102 [122-85]	115 [167-79]
BUBR1	Chip 4	β-actin & GAPDH	100 [119-84]	106 [130-86]	76 [89-65]	77 [99-60]
MAD1	Chip 5	β-actin & GAPDH	100 [147-68]	54 [67-43]	122 [158-94]	68 [92-50]
MAD2L1	Chip 5	β-actin & GAPDH	100 [139-72]	97 [122-77]	126 [184-86]	115 [137-96]
MAD2L2	Chip 5	β-actin & GAPDH	100 [128-78]	61 [71-52]	93 [114-76]	57 [64-51]
CHK1	Chip 5	β-actin & GAPDH	100 [131-76]	78 [94-65]	72 [85-61]	88 [103-75]
CHK2	Chip 5	β-actin & GAPDH	100 [156-64]	82 [103-65]	93 [100-86]	113 [162-79]
BUB1	Chip 4	β-actin & GAPDH	100 [119-84]	82 [107-63]	59 [81-43]	76 [92-63]
BUB3	Chip 4	β-actin & GAPDH	100 [122-82]	79 [100-63]	71 [88-58]	59 [70-50]
ATM	Chip 4	β-actin & GAPDH	100 [127-79]	115 [148-90]	87 [123-61]	118 [156-89]
FANCD2	Chip 5	β-actin & GAPDH	100 [142-70]	76 [82-71]	81 [94-70]	72 [80-65]
PARP1	Chip 5	β-actin & GAPDH	100 [141-71]	70 [82-59]	79 [93-67]	79 [96-65]
RAD51	Chip 5	β-actin & GAPDH	100 [150-67]	96 [105-88]	95 [118-77]	89 [113-70]
RAD51C	Chip 5	β-actin & GAPDH	100 [158-63]	73 [107-50]	94 [116-76]	102 [145-71]
RAD51D	Chip 5	β-actin & GAPDH	100 [129-78]	42 [50-35]	79 [86-72]	49 [64-38]
BARD1	Chip 5	β-actin & GAPDH	100 [156-64]	117 [126-108]	240 [279-207]	81 [90-73]
BRCA1	Chip 4	β-actin & GAPDH	100 [127-79]	101 [123-83]	68 [83-56]	89 [111-71]
BRCA2	Chip 4	β-actin & GAPDH	100 [121-82]	89 [110-72]	83 [104-66]	90 [120-67]
MERIT40	Chip 2	β-actin	100 [121-83]	104 [130-83]	44 [56-35]	66 [74-59]
BRE	Chip 4	β-actin & GAPDH	100 [125-80]	106 [145-77]	81 [103-64]	100 [138-72]
CCDC98	Chip 4	β-actin & GAPDH	100 [155-65]	110 [139-87]	59 [73-48]	87 [106-72]
RAP80	Chip 5	β-actin & GAPDH	100 [129-78]	60 [66-54]	75 [100-56]	60 [76-48]
BRCC3	Chip 1	β-actin & GAPDH	100 [122-82]	125 [154-102]	88 [100-77]	121 [160-91]
ERCC1	Chip 5	β-actin & GAPDH	100 [145-69]	55 [75-41]	117 [126-108]	66 [99-44]
ERCC2	Chip 5	β-actin & GAPDH	100 [130-77]	101 [114-90]	75 [84-67]	140 [162-121]
XRCC5	Chip 5	β-actin & GAPDH	100 [134-74]	93 [113-76]	105 [131-84]	104 [125-87]
XRCC6	Chip 5	β-actin & GAPDH	100 [133-75]	63 [73-54]	73 [84-63]	63 [76-52]
POLB	Chip 5	β-actin & GAPDH	100 [133-75]	69 [78-61]	128 [141-117]	56 [67-47]
MLH1	Chip 5	β-actin & GAPDH	FAILED	FAILED	FAILED	FAILED
MSH2	Chip 5	β-actin & GAPDH	100 [145-69]	56 [78-40]	83 [108-64]	68 [82-56]
PMS1	Chip 5	β-actin & GAPDH	100 [137-73]	77 [106-56]	78 [98-62]	85 [103-70]
PIK3CA	Chip 5	β-actin & GAPDH	100 [133-75]	47 [63-35]	61 [75-50]	50 [59-42]
KRAS	Chip 5	β-actin & GAPDH	100 [136-73]	94 [106-84]	82 [96-70]	105 [126-87]
BRAF	Chip 4	β-actin & GAPDH	100 [128-78]	106 [158-71]	80 [116-55]	83 [127-54]
CDK12	Chip 6	β-actin & GAPDH	100 [121-82]	122 [160-93]	103 [115-93]	90 [94-86]
TP53	Chip 5	β-actin & GAPDH	100 [135-74]	64 [71-57]	93 [115-75]	75 [90-62]
ABHD8	Chip 2	β-actin	100 [126-79]	46 [63-33]	104 [140-77]	59 [69-50]
DDA1	Chip 2	β-actin	100 [128-78]	53 [67-42]	108 [141-82]	79 [98-64]
GTPBP3	Chip 2	β-actin	100 [122-82]	59 [73-47]	121 [149-98]	79 [91-69]
MRPL34	Chip 3	β-actin & GAPDH	100 [125-80]	36 [54-24]	91 [106-78]	73 [89-60]
MYO9B	Chip 3	β-actin & GAPDH	100 [129-78]	90 [117-69]	82 [98-69]	76 [83-69]
NR2F6	Chip 3	β-actin & GAPDH	100 [128-78]	55 [72-42]	114 [143-91]	79 [89-70]
USE1	Chip 4	β-actin & GAPDH	100 [119-84]	88 [124-62]	69 [79-60]	67 [101-45]
β-actin SET 1	Chip 1	GAPDH	100 [131-76]	79 [104-60]	103 [126-85]	102 [131-79]
GAPDH SET 1	Chip 1	β-actin	100 [131-76]	126 [166-96]	97 [118-80]	98 [126-76]
GAPDH SET 2	Chip 2	N/A (β-actin failed)	N/A	N/A	N/A	N/A
β-actin SET 3	Chip 3	GAPDH	100 [138-73]	92 [132-64]	123 [157-96]	143 [168-122]
GAPDH SET 3	Chip 3	β-actin	100 [138-73]	108 [155-75]	81 [103-63]	70 [82-60]
β-actin SET 4	Chip 4	GAPDH	100 [127-79]	81 [107-61]	153 [186-126]	130 [162-104]
GAPDH SET 4	Chip 4	β-actin	100 [127-79]	127 [161-100]	65 [79-53]	77 [96-62]
β-actin SET 5	Chip 5	GAPDH	100 [144-69]	116 [148-91]	124 [147-105]	100 [130-77]
GAPDH SET 5	Chip 5	β-actin	100 [144-69]	86 [110-68]	81 [96-69]	100 [129-77]
β-actin SET 6	Chip 6	GAPDH	100 [123-81]	86 [109-68]	128 [149-110]	135 [161-113]
GAPDH SET 6	Chip 6	β-actin	100 [123-81]	117 [149-92]	78 [91-67]	74 [88-62]
GAPDH SET 7	Chip 7	N/A (β-actin failed)	N/A	N/A	N/A	N/A